



SEQUENCE LISTING

<110> Barnett, Susan  
Zur Megede, Jan

<120> POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

<130> PP01631.101

<140> 09/475,704  
<141> 1999-12-30

<150> 09/610,313  
<151> 2000-07-05

<160> 46

<170> PatentIn Ver. 2.0

<210> 1  
<211> 60  
<212> DNA  
<213> Human immunodeficiency virus

<400> 1  
gacatcaagc agggcccaa ggagccttc cgcgactacg tggaccgctt cttcaagacc 60

<210> 2  
<211> 60  
<212> DNA  
<213> Human immunodeficiency virus

<400> 2  
gacatccgcc agggcccaa ggagccttc cgcgactacg tggaccgctt cttcaagacc 60

<210> 3  
<211> 1479  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic Gag  
of HIV strain AF110965

<400> 3  
atgggcgccc gcccagcat cctgcgcggc ggcaagctgg acgcctggga gcgcattccgc 60  
ctgcgcggcg gcggcaagaa gtgtacatg atgaagcacc tgggtggc cagccgcgag 120  
ctggagaagt tcgcccgtaa cccccggctg ctggagacca gcgagggctg caagcagatc 180  
atccgcgcgc tgacccccgc cctgcagacc ggcagcgagg agctgaagag cctgttcaac 240  
accgtggcca ccctgtactg cgtgcacgag aagatcgagg tccgcgacac caaggaggcc 300  
ctggacaaga tcgaggagga gcagaacaag tgccagcaga agatccagca ggccgaggcc 360  
gccgacaagg gcaagggtgag ccagaactac cccatcgtgc agaacctgca gggccagatg 420  
gtgcaccagg ccatcagccc ccgcacctg aacgcctggg tgaagggtgat cgaggagaag 480  
gccttcagcc ccgagggtgat cccatgttc accgcctgaa gcgaggggcgc caccggcc 540  
gacctgaaca cgatgtgaa caccgtggc ggccaccagg cccatgca gatgctgaag 600  
gacaccatca acgaggaggc cggcagtgac accccgtgca cgccggcccc 660

atcgcccccgc gccagatgcg cgagccccgc ggcagcgaca tcgcccggcac caccaggcacc 720  
ctgcaggagc agatcgccctg gatgaccagc aaccccccga tccccgtggg cgacatctac 780  
aagcggtgga tcatcctggg cctgaacaag atcgtgcggta tgtacagccc cgtgagcatc 840  
ctggacatca agcagggccc caaggagccc ttccgcgact acgtggaccg cttcttcaag 900  
accctgcgcg ccgagcagag caccaggag gtgaagaact ggatgaccga caccctgctg 960  
gtgcagaacg ccaaccccgaa ctgcaagagacc atcctgcgcg ctctcgccccc cggcgccagc 1020  
ctggaggaga tgatgaccgc ctgcccaggc gtggcgccccc ccagccacaa ggcccgcgtg 1080  
ctggccgagg cgatgagcca ggcacacacc agcgtgatga tgcagaagag caacttcaag 1140  
ggcccccggc gcatcgtaa gtgcttcaac tgcggcaagg agggccacat cgcccccaac 1200  
tgccgcgccc cccgcaagaa ggctgctgg aagtgcggca aggagggcca ccagatgaag 1260  
gactgcaccc agcgccaggc caacttcctg ggcaagatct ggcccagcca caagggccgc 1320  
cccgcaact tcctgcagag ccgccccgag cccaccgccc ccccccggc gagcttccgc 1380  
ttcgaggaga ccaccccccgg ccagaaggcag gagagcaagg accgcgagac cctgaccagc 1440  
ctgaagagcc tgttcggcaa cgacccctg agccagtaa 1479

<210> 4  
<211> 1509  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic Gag  
of HIV strain AF110967

<400> 4  
atgggcggcc gcgcacgcat cctgcgcggc gagaagctgg acaagtggga gaagatccgc 60  
ctgcgcggcc gcgcaagaa gcactacatg ctgaagcacc tggtgtgggc cagccgcgag 120  
ctggagggtc tcgcccctgaa ccccgccctg ctggagaccg cccggggctg caagcagatc 180  
atgaagcagc tgcagccgc cctgcagacc ggcacccgagg agctgcgcag cctgtacaac 240  
accgtggcca ccctgtactg cgtgcacgccc ggcacatcgagg tccgcgcacac caaggaggcc 300  
ctggacaaga tcgaggagga gcagaacaag tcccagcaga agaccccgca ggcacaggag 360  
gccgacggca aggtgagcca gaactacccc atcgtgcaga acctgcaggcc ctagatggtg 420  
caccaggcca tcagcccccg caccctgaac gcctgggtga aggtgatcga ggagaaggcc 480  
ttcagccccg aggtgatccc catgttcacc gccctgagcg agggcgccac ccccccaggac 540  
ctgaacacga tggtaaacac cgtggcgcc caccaggccc ccatgcagat gctgaaggac 600  
accatcaacg aggaggccgc cgagtggac cgcctgcacc cctgcaggcc cggccccgtg 660  
gccccccggcc agatgcgcga ccccccggc acgcacatcg cccggccac cagcaccctg 720  
caggagcaga tcgcctggat gaccagcaac ccccccgtgc ccgtggcgca catctacaag 780  
cggtggatca tcctgggcct gaacaagatc gtgcggatgt acagccccgt ggcacccctg 840  
gacatccgcg aggcccccaa ggagcccttc cgcgactacg tggaccgctt cttcaagacc 900  
ctgcgcgcgc agcaggccac ccaggacgtg aagaacttggta tgaccgagac cctgctggtg 960  
cagaacgcgc accccgactg caagaccatc ctgcgcgcctc tcggccccgg cgccacccctg 1020  
gaggagatga tgaccgcctg ccaggcggtg ggcggccccg gccacaaggc cccgcgtctg 1080  
gccgaggcgca tgagccaggc caacagcggtg aacatcatga tgcagaagag caacttcaag 1140  
ggccccccggc gcaacgtcaa gtgcttcaac tgcggcaagg agggccacat cgccaaagaac 1200  
tgccgcgcgc cccgcaagaa ggctgctgg aagtgcggca aggagggcca ccagatgaag 1260  
gactgcaccc agcgccaggc caacttcctg ggcaagatct ggcccagcca caagggccgc 1320  
cccgcaact tcctgcagaa ccgcagcgag cccgcgcggc ccaccgtgcc caccggcccc 1380  
cccgccgaga gcttccgcctt cgaggagacc acccccccggc ccaaggcagga gcccacaggac 1440  
cgcgagccct accgcgagcc cctgaccgcctc ctgcgcagcc tgttcggcag cggccccctg 1500  
agccagtaa 1509

<210> 5  
<211> 141  
<212> DNA  
<213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Env common  
     region of HIV strain AF110968

<400> 5  
 accatcacca tcacacctggc catcaagcgat atcatcaaca tgtggcagaa ggtggccgc 60  
 gccatgtacg cccccccat cgccggcaac ctgacacctgc agagcaacat caccggcctg 120  
 ctgctgaccc gcgacggcgg c 141

<210> 6  
 <211> 1431  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic  
     gp120 coding region of HIV strain AF110968

<400> 6  
 agcgtggtgg gcaacctgtg ggtgaccgtg tactacggcg tgcccgtgtg gaaggaggcc 60  
 aagaccaccc tggctgcac cagcgacgcc aaggcctacg agaccgaggt gcacaacgtg 120  
 tggccaccc acgcctgcgt gcccaccgc cccaaccccc aggagatcgt gctggagaac 180  
 gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240  
 atcagcctgt gggaccagag cctgaagccc tgcgtgaagc tgaccccccgt gtcgtgacc 300  
 ctgaagtgc gcaacgtgaa cgccaccaac aacatcaaca gcatgatgca caacagcaac 360  
 aaggcgaga tgaagaactg cagttcaac gtgaccaccc agctgcgcga ccgcaagcag 420  
 gaggtgcacg ccctgttcta ccgcctggac gtggtgcccc tgcaggcCAA caacagcaac 480  
 gagtaccgcg tgcgttcta ccgcctggac gtggtgcccc tgcaggcCAA caacagcaac 540  
 ttcgacccca tcccccattca ctactgcacc cccgcccgc acgcacatccgt gaagtgcac 600  
 aaccagaccc tcaacggcac cgcccccgc aacaacgtgaa gcatgatgca gtgcgcac 660  
 ggcataaggc ccgtggtag caccagctg ctgctgaaacg gcatgatgca acgcacatccgt 720  
 atcatcatcc gcaaggcgagaa cctggccaaac aacgccaaga tcatcatcg gcatgatgca 780  
 aagcccgta agatcggtg cgtgcgcaccc aacaacaaca cccgcaagag cgtgcgcac 840  
 ggccccggcc agaccttcta ccgcctggc gagatcatcg gcatgatgca acgcacatccgt 900  
 tgcatacatca acaagaccga gtggaacacg accctgcagg gcatgatgca acgcacatccgt 960  
 gagacttca gcaagaaggc catcaagttc gagcccgac gcatgatgca acgcacatccgt 1020  
 accacccaca gcttcaactg ccgcggcgag ttcttctact gcatgatgca acgcacatccgt 1080  
 aacagcaccc acagccccag cttcaacggc accgagaaca gcatgatgca acgcacatccgt 1140  
 atcacctgc gcatcaagca gatcatcaac atgtggcaga aggtggccgc gcatgatgca acgcacatccgt 1200  
 gccccccca tcgcccggcaa cctgacacctgc gagagcaaca tcacccggcct gcatgatgca acgcacatccgt 1260  
 cgcacggcg gcaagaccgg ccccaacgc accgagatct tccggcccccgg gcatgatgca acgcacatccgt 1320  
 atgcgcgaca actggcgcaaa cggactgtac aagtacaagg tggtgagat gcatgatgca acgcacatccgt 1380  
 ggcgtggccc ccaccggaggc caagcgccgc gtggtgagc gcatgatgca acgcacatccgt 1431

<210> 7  
 <211> 1944  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic  
     gp140 coding region of HIV strain AF110968

<400> 7  
 agcgtggtgg gcaacctgtg ggtgaccgtg tactacggcg tgcccgtgtg gaaggaggcc 60  
 aagaccaccc tggctgcac cagcgacgcc aaggcctacg agaccgaggt gcacaacgtg 120

tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcgt gctggagaac 180  
 gtgaccgaga acttcaacat gtggaagaac gacatggtg accagatgca cgaggacatc 240  
 atcagcctgt gggaccagag cctgaagccc tgcgtgaagc tgaccccccgt gtcgtgacc 300  
 ctgaagtgcc gcaacgtgaa cggcaccaac aacatcaaca gcatgatgca caacagcaac 360  
 aagggcgaga tgaagaactg cagcttcaac gtgaccaccg agctgcgcga ccgcaagcag 420  
 gaggtgcacg ccctgttcta cgcctggac gtggtcccc tgcaggcga caacagcaac 480  
 gagtacctcc tcatcaactg caacaccgac gccatcaccc aggctgccc caaggtgagc 540  
 ttcgacccca tccccatcca ctactgcacc cccgcggct acgccatcct gaagtgcaac 600  
 aaccagaccc tcaacgcac cgccccctgc aacaacgtga gcagcgtgca gtgcgcac 660  
 ggcataaggc ccgtggtag caccctgctg ctgctgaacg gcagcctggc caagggcgag 720  
 atcatcatcc gcagcgagaa cctggccaac aacgccaaga tcatcatcgt gcagctgaaac 780  
 aagcccgta agatcgtgtg cgtgcgcccc aacaacaaca cccgcaagag cgtgcgcac 840  
 ggccccggcc agaccattcta cgcaccggc gagatcatcg gcgacatccg ccaggcctac 900  
 tgcattatca acaagaccga gtggaacacgc accctgcagg gcgtgagcaa gaagctggag 960  
 gagcaattca gcaagaaggc catcaagttc gagccagca gcggcggcga cctggagatc 1020  
 accacccaca gcttcaactg ccgcggcga ttcttctact gcgacaccag ccagctgttc 1080  
 aacagcaccc acagccccag ctcaacggc accgagaaca agctgaacgg caccatcacc 1140  
 atcacctgcc gcatcaagca gatcatcaac atgtggcaga aggtgggccc gcgcattgtac 1200  
 gccccccca tcgcccggaa cctgacctgc gagagcaaca tcaccggcct gtcgtgacc 1260  
 cgcgcggcga gcaagaccgg ccccaacgcac accgagatct tccggccccc cggcggcga 1320  
 atgcgcgaca actggcgcac cgagctgtac aagtacaagg ttgtggagat caagccctg 1380  
 ggcgtggccc ccaccggaggc caagcgccgc gtggtggagc gcgagaagcg cgccgtggc 1440  
 atcgccgcgg tggctctggg ctgcctggc gcccggca gcaccatggg cgccgcggc 1500  
 atcaccctga ccgtgcaggc ccgcctgtc ctgagcggca tcgtgcagca gcagaacaac 1560  
 ctgctgcgcg ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 1620  
 cagctgcaga cccgcattct gggcgtggag cgctacctga aggaccagca gtcgtggc 1680  
 atctggggct gcagcgcaaa gctgatctgc accaccggc tgccctggaa cagcagctgg 1740  
 agcaaccgca gccacgacga gatctgggac aacatgaccc ggatgcagtg ggaccggcag 1800  
 atcaacaact acaccgacac catctaccgc ctgctggagg agagccagaa ccagcaggag 1860  
 aagaacgaga aggacctgct gcccctggac agctggcaga acctgtggaa ctgggtcagc 1920  
 atcaccact ggctgtggta catc 1944

<210> 8  
 <211> 2466  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic  
 gp160 coding region of HIV strain AF110968

<400> 8  
 agctgtgttgg gcaaccctgtg ggtgaccgtg tactacggcg tgccctgtg gaaggaggcc 60  
 aagaccaccc tggctctgcac cagcgacgc aaggctacg agaccgagggt gcacaacgtg 120  
 tggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcgt gctggagaac 180  
 gtgaccgaga acttcaacat gtggaagaac gacatggtg accagatgca cgaggacatc 240  
 atcagcctgt gggaccagag cctgaagccc tgcgtgaagc tgaccccccgt gtcgtgacc 300  
 ctgaagtgcc gcaacgtgaa cggcaccaac aacatcaaca gcatgatgca caacagcaac 360  
 aagggcgaga tgaagaactg cagcttcaac gtgaccaccg agctgcgcga ccgcaagcag 420  
 gaggtgcacg ccctgttcta cgcctggac gtggtcccc tgcaggcga caacagcaac 480  
 gagtacctcc tccccatcca ctactgcacc cccgcggct acgccatcct gaagtgcaac 540  
 aaccagaccc tcaacgcac cgccccctgc aacaacgtga gcagcgtgca gtgcgcac 600  
 ggcataaggc ccgtggtag caccctgctg ctgctgaacg gcagcctggc caagggcgag 660  
 atcatcatcc gcagcgagaa cctggccaac aacgccaaga tcatcatcgt gcagctgaaac 720  
 aagcccgta agatcgtgtg cgtgcgcccc aacaacaaca cccgcaagag cgtgcgcac 780  
 ggccccggcc agaccattcta cgcaccggc gagatcatcg gcgacatccg ccaggcctac 900

tgcatcatca	acaagaccga	gtgaaacaggc	accctgcagg	gcgtgagcaa	gaagctggag	960
gagacttca	gcaagaaggc	catcaagtgc	gagcccgagca	gcggcggcga	cctggagatc	1020
accacccaca	gcttcaactg	ccgcggcgag	ttcttctact	gcgacaccag	ccagctgttc	1080
aacagcacct	acagccccag	cttcaacggc	accgagaaca	agctgaacgg	caccatcacc	1140
atcacctgcc	gcatcaagca	gatcatcaac	atgtggcaga	aggtgggccc	cgccatgtac	1200
gcccccccca	tcgcccggcaa	cctgacctgc	gagagcaaca	tcaccggct	gctgctgacc	1260
cgcgacggcg	gcaagaccgg	ccccaaacgac	accgagatct	tccgccccgg	cgccggcgcac	1320
atgcgcgaca	actggcgcaa	cgagctgtac	aagtacaagg	ttgtggagat	caagccccctg	1380
ggcgtggccc	ccaccggaggc	caagcggcgc	gtggtgagc	gcgagaagcg	cgccgtggc	1440
atccggcgccg	tgttcctggg	cttcctggc	gccgcccggca	gcaccatggg	cgccggccagc	1500
atcacccctga	ccgtgcaggc	ccgcctgctg	ctgagcggca	tcgtgcagca	gcagaacaac	1560
ctgctgcgcg	ccatcgaggc	ccagcagcac	ctgctgcagc	tgaccgtgt	ggccatcaag	1620
cagctgcaga	cccgcatcct	ggccgtggag	cgctacctga	aggaccagca	gctgctggc	1680
atctggggct	gcagcggcaa	gctgatctgc	accaccggc	tgcctggaa	cagcagctgg	1740
agcaaccgca	gccacgacga	gatctggac	aacatgacct	ggatgcagtg	ggaccgcgag	1800
atcaacaact	acaccgacac	catctaccgc	ctgctggagg	agagccagaa	ccagcaggag	1860
aagaacgaga	aggacctgct	ggccctggac	agctggcaga	acctgtggaa	ctgggtcagc	1920
atcaccaact	ggctgtggta	catcaagatc	ttcatcatga	tcgtggccgg	cctgatcggc	1980
ctgcgcata	tcttcggcgt	gctgagcatc	gtgaaccgcg	tgccgcaggg	ctacagcccc	2040
ctgcccattc	agaccctgac	ccccaaacccc	cgcgagcccg	accgcctggg	ccgcatcgag	2100
gaggagggcg	gcgagcagga	ccgcggccgc	agcatccgc	ttgtgagcgg	ttccctggcc	2160
ctggcctggg	acgacctgct	cagcctgtgc	ctgttcagct	accaccgcct	gcgcgacttc	2220
atccgtatcg	ccgcccgcgt	gctggagctg	ctggccagc	gcccgtggg	ggccctgaag	2280
tacctgggca	gcctggtgca	gtactggggc	ctggagctga	agaagagcgc	catcagcctg	2340
ctggacacca	tcgccatcgc	cgtggccgag	ggcaccgacc	gcatcatcga	gttcatccag	2400
cgcacatctg	gcgcacatccg	caacatcccc	cgccgcacatcc	gccaggcgtt	cgaggccgccc	2460
ctgcag						2466

<210> 9  
<211> 2547  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic  
signal sequence and gp160 coding region of HIV  
strain AF110968

<400> 9  
atgcgcgtga tgggcacatcct gaagaactac cagcagtggt ggatgtgggg catcctggc 60  
ttctggatgc tgatcatcag cagcgtggtg ggcaacctgt gggtgaccgt gtactacggc 120  
gtgcccgtgt ggaaggaggc caagaccacc ctgttctgca ccagcgcacgc caaggcctac 180  
gagaccgagg tgcacaacgt gtggccacc cacgcctgcg tgcccacca ccccaacccc 240  
caggagatcg tgctggagaa cgtgaccgag aacttcaaca tgtgaaagaa cgacatggtg 300  
gaccagatgc acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtgaag 360  
ctgacccccc tgcgttgac cctgaagtgc cgcaacgtga acgcccaccaa caacatcaac 420  
agcatgtatcg acaacagcaa caaggcgag atgaagaact gcagcttcaa cgtgaccacc 480  
gagctgcgcg accgcaagca ggaggtgcac gccctgttct accgcctggc cgtggtgcgg 540  
ctgcaggcga acaacagcaa cgagtaccgc ctgatcaact gcaacaccag cgccatcacc 600  
caggcctgccc ccaagggtgag cttcgaccccc atccccatcc actactgcac ccccgccggc 660  
tacccatcc tgaagtgc aaaccagacc ttcaacggca cccggccctg caacaacgtg 720  
agcagcgtgc agtgcgccta cgcatcaag cccgtggtga gcacccagct gctgctgaac 780  
ggcagcctgg ccaaggcgaa gatcatcatc cgcagcgaga acctggccaa caacgccaag 840  
atcatcatcg tgcagctgaa caagccctgt aagatcgtgt gcgtgcgcgg caacaacaac 900  
acccgcaaga gcgtgcgcata cggccccggc cagacccttct acgcaccagg cgagatcatc 960  
ggcgcacatcc gccaggccta ctgcacatc aacaagaccg agtggaaacag caccctgcag 1020  
ggcgtgagca agaagctgga ggagcacttc agcaagaagg ccatcaagtt cgagcccgagc 1080

agcggcggcg	acctggagat	caccacccac	agcttcaact	gccgcggcga	gttcttctac	1140
tgcacacca	gccagctgtt	caacagcacc	tacagcccc	gcttcaacgg	caccgagaac	1200
aagtgaacg	gcaccatcac	catcacctgc	cgcataaggc	agatcatcaa	catgtggcag	1260
aagtgggcc	gcccattgt	cgccccccc	atgcggcga	acctgacctg	cgagagcaac	1320
atcaccggcc	tgctgtgac	ccgcgacggc	ggcaagaccg	gccccaaacga	caccgagatc	1380
ttccggcccg	gcggcggcga	catgcgcgac	aactggcga	acgagctgt	caagtacaag	1440
gtgtggaga	tcaagccct	gggcgtggcc	cccaccgagg	ccaaggcggc	cgtggtgag	1500
cgcagaagc	gcgcgtggg	catcgccgc	gtttcctgg	gcttcctggg	cgccgcggc	1560
agcaccatgg	gcgcgcggc	catcacctg	accgtgcagg	cccgcctgt	gctgagcggc	1620
atcgctcagc	agcagaacaa	cctgctgcgc	gccatcgagg	cccagcagca	cctgctgcag	1680
ctgaccgtgt	ggggcatcaa	gcagctgcag	accgcaccc	tggccgtgga	gctgacccctg	1740
aaggaccaggc	agctgtggg	catctggggc	tgcagcggca	agctgatctg	caccacccgc	1800
gtgccctgga	acagcagctg	gagcaaccgc	agccacgcac	agatctggg	caacatgacc	1860
tggatgcagt	gggaccgcga	gatcaacaac	tacaccgaca	ccatctaccg	cctgctggag	1920
gagagccaga	accagcagga	gaagaaccgag	aaggacctgc	tggccctgga	cagctggcag	1980
aacctgtgga	actggttcag	catcaccaac	tggctgttgt	acatcaagat	tttcatcatg	2040
atcggtggcg	gcctgatcgg	cctgcgcatac	atttcgccg	tgctgagcat	cgtgaaccgc	2100
gtgcgcctagg	gctacagccc	cctgccttc	cagacccctga	cccccaaccc	ccgcgagccc	2160
gaccgcctgg	gcgcatcga	ggaggagggc	ggcgagcagg	accgcggccg	cagcatccgc	2220
ctgtgtgagcg	gcttcctggc	cctggcctgg	gacgacctgc	gcagcctgt	cctgttcagc	2280
taccaccgccc	tgcgcgactt	catcctgatc	gccgcggcgc	tgctggagct	gctggccag	2340
cgccgctggg	aggccctgaa	gtacctgggc	agcctggtc	agtaactgggg	cctggagctg	2400
aagaagagcg	ccatcagcct	gctggacacc	atgcacatcg	ccgtggccga	gggcacccgac	2460
cgcacatcatcg	agttcatcca	gcccacatcg	cgccatcc	gcaacatccc	ccgcgcacatc	2520
cgccaggggct	tgcaggccgc	cctgcag				2547

<210> 10  
<211> 1035  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic a  
gp41 coding region of HIV strain AF110968

<400> 10						
gcccgtggca	tcggcggccgt	gttcctgggc	ttcctggcg	ccgcggcag	caccatggc	60
gcccgcagca	tcaccctgac	cgtgcaggcc	cgcctgctgc	tgagcggcat	cgtgcagcag	120
cagaacaacc	tgctgcgcgc	catcgaggcc	cagcagcacc	tgctgcagct	gaccgtgtgg	180
ggcatcaagc	agctgcagac	ccgcacatctg	gccgtggagc	gctacctgaa	ggaccagcag	240
ctgtgtggca	tctgggctg	cagcggcaag	ctgatctgc	ccaccgcgt	gccctggAAC	300
agcagctgga	gcaaccgcag	ccacgcacgag	atctggaca	acatgacctg	gatgcagtgg	360
gaccgcgaga	tcaacaacta	caccgacacc	atctaccgc	tgctggagga	gagccagaac	420
cacaggaga	agaacgagaa	ggacctgtg	gccctggaca	gctggcagaa	cctgtggAAC	480
tggttcagca	tcaccaactg	gctgtggtac	atcaagatct	tcatcatgat	cgtggcggc	540
ctgatcgcc	tgcgcacat	cttcgcgt	ctgagacatcg	tgaaccgcgt	gcccaggggc	600
tacagccccc	tgcccttcca	gaccctgacc	cccaacccccc	gcgagcccg	ccgcctggc	660
cgcacatcgagg	aggaggcgg	cgacgcggac	cgcggccgc	gatccgcct	ggtgagcggc	720
ttcctggccc	ttggcctggg	cgacctgcgc	agcctgtgcc	tgttcagcta	ccaccgcctg	780
cgcacttca	tcctgatcgc	cgcgcgcgt	ctggagctgc	tggccagcg	cggctggag	840
gccctgaagt	acctggcag	cctgggtgcag	tactggggcc	tggagctgaa	gaagagcggc	900
atcagcctgc	tggacacccat	cgccatcgcc	gtggccgagg	gcaccgaccc	catcatcgag	960
ttcatccagc	gcatctgcgc	cgccatccgc	aacatcccc	gccgcacatcc	ccagggtttc	1020
gaggccgccc	tgcag					1035

<210> 11  
<211> 144

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic Env  
common region of HIV strain AF110975

<400> 11  
agcatcatca ccctgcctg ccgcataaag cagatcatcg acatgtggca gaagggtggc 60  
cgcgcattt acgccccccc catcgaggc aacatcacct gcagcagcag catcaccggc 120  
ctgctgctgg cccgcgacgg cgcc 144

<210> 12  
<211> 1437  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic  
gp120 coding region of HIV strain AF110975

<400> 12  
agccgcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60  
agcaccaccc ttttcgcgc cagcgcacgc aaggcctacg agaaggaggt gcacaacgtg 120  
tggccaccc acgcctgcgt gcccaccgc cccaaacccc aggagatcga gctggacaac 180  
gtgaccgaga acttcaacat gtggaaaac gacatggtgg accagatcga cgaggacatc 240  
atcagcctgt gggaccagag cctgaagccc cgcgtgaagc tgaccccccgtgtgcgtgacc 300  
ctgaagtgc ccaactacag caccacactac agcaacacca tgaacgccac cagctacaac 360  
aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccga gctgcgcgac 420  
aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgccttgaacac 480  
agcagcagtg accgcctgtat caactgcac accagcgcca tcacccaggc ctgccccaaag 540  
gtgagcttcg accccatcccc catccactac tgcgccttgc cccgctacgc catcctgaag 600  
tgcaagaaca acaccagcaa cgccacccgc ccctggcaga acgtgagcac cgtgcagtgc 660  
accacacggca tcaagcccgt ggtgagcacc cccctgtgc tgaacggcag cctggccgag 720  
ggccgcgaga tcatcatccg cagcaagaac ctgagacaaca acgcctacac catcatcgtg 780  
cactgaacg acagcgtgga gatcgtgtgc acccgccccca acaacaacac cccgcaaggcc 840  
atccgcacatcg gccccggcca gaccttctac gccaccgaga acatcatcg cgacatccgc 900  
cagccccact gcaacatcag cggccggcgag tggaaacaagg ccgtgcagcg cgtgagcgcc 960  
aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccgccagc cggccggcgac 1020  
ctggagatca ccacccacag ctcaactgc cgcggcgagt tcttctactg caacaccagc 1080  
aagctgttca acagcagcta caacggcacc agtaccgcg gcaccgagag caacagcagc 1140  
atcatcaccc tggccctgcgc catcaagcag atcatcgaca tgtggcagaa ggtggccgc 1200  
gccccatcg ccccccacat cgaggccaaat atcaccgtca gcagcagcat caccggcctg 1260  
ctgctggccc gcgacggcg cctggacaac atcaccaccg agatcttccg ccccccaggcc 1320  
ggcgcacatga aggacaactg ggcacacgag ctgtacaagt acaagggtggt ggagatcaag 1380  
ccctggcg tggcccccac cgaggccaaag cggccgtgg tggagcgcga gaagcgc 1437

<210> 13  
<211> 1950  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic  
gp140 coding region of HIV strain AF110975

<400> 13

agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60  
 agcaccaccc ttttctgcgc cagcgcacgaa aaggccatcg agaaggaggt gcacaacgtg 120  
 tggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcg aactggacaac 180  
 gtgaccgaga acttcaacat gtggaagaac gacatggtg accagatgca cgaggacatc 240  
 atcagcctgt gggaccagag cctgaagccc cgctgtaaagc tgaccccccgt gtgcgtgacc 300  
 ctgaagtgc acaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360  
 aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccgac gctgcgcac 420  
 aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgcccct gaacagcaac 480  
 agcagcgagt accgcctgtat caactgcaac accagcgcc tcacccaggc ctgccccaaag 540  
 gtgagcttcg accccatccc catccactac tgcgcggcc cccgctacgc catcctgaag 600  
 tgcaagaaca acaccagcaa cgccaccggc ccctgcccaga acgtgagcac cgtcagtg 660  
 acccacggca tcaagccgt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720  
 ggccgcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcg 780  
 cactgaacg acagcgtgga gatcgtgtgc acccgccca acaacaacac ccgcaaggcc 840  
 atccgcatcg gccccggcca gaccttctac gccaccgaga acatcatcg cgacatccgc 900  
 cagccccact gcaacatcag cgccggcgag tggacaagg ccgtgcagcg cgtgagcgcc 960  
 aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccagcag cggccgcac 1020  
 ctggagatca ccacccacag cttaactgc cgccggcgagt tcttctactg caacaccagc 1080  
 aagctgttca acagcagcta caacggcacc agtaccgcg gcaccgagag caacagcagc 1140  
 atcatcaccc tggccctgccc catcaagcag atcatcgaca tgtggcagaa ggtggccgc 1200  
 gccatctacg ccccccattt cgaggcaac atcacctgca gcagcagcat caccggcctg 1260  
 ctgtggccc gcgacggcg cctggacaac atcaccaccg agatcttccg ccccccaggcc 1320  
 ggcacatga aggacaactg ggcacacgag ctgtacaagt acaagggtggt ggagatcaag 1380  
 ccctggcg tggcccccac cgaggcaag cgccggcg tggagcgcga gaagcgcgc 1440  
 gtggcatcg ggcgcgtat cttcggttcc ctggccggcc cccggcagcaa catggccgc 1500  
 gccagcatca ccctgaccgc ccaggcccgc cagctgctga gccgcacatcg gcagcagcag 1560  
 agcaacctgc tgcgcgcctt cgaggcccag cagcacatgc tgcagctgac cgtgtgggc 1620  
 atcaagcagc tgcaggcccg ctgtctggcc atcgagcgct acctgaagga ccagcagctg 1680  
 ctggcatct ggggctgcag cgcaagctg atctgcacca ccaccgtgcc ctggAACAGC 1740  
 agctggagca acaagacca gggcagatc tggagaaca tgacctggat gcagtggac 1800  
 aagagatca gcaactacac cgcatcatc taccgcctgc tggaggagag ccagaaccag 1860  
 cagggcaga acgagaagga cctgctggcc ctggacagcc gcaacaacct gtggagctgg 1920  
 ttcaacatca gcaactggct gtggatcatc 1950

<210> 14

<211> 2493

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic  
gp160 coding region of HIV strain AF110975

<400> 14

agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60  
 agcaccaccc ttttctgcgc cagcgcacgaa aaggccatcg agaaggaggt gcacaacgtg 120  
 tggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcg aactggacaac 180  
 gtgaccgaga acttcaacat gtggaagaac gacatggtg accagatgca cgaggacatc 240  
 atcagcctgt gggaccagag cctgaagccc cgctgtaaagc tgaccccccgt gtgcgtgacc 300  
 ctgaagtgc acaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360  
 aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccgac gctgcgcac 420  
 aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgcccct gaacagcaac 480  
 agcagcgagt accgcctgtat caactgcaac accagcgcc tcacccaggc ctgccccaaag 540  
 gtgagcttcg accccatccc catccactac tgcgcggcc cccgctacgc catcctgaag 600  
 tgcaagaaca acaccagcaa cgccaccggc ccctgcccaga acgtgagcac cgtcagtg 660  
 acccacggca tcaagccgt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720

ggccggcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcgtg 780  
 cacctgaacg acagcgtgga gatcgtgtgc acccgccccca acaacaacac ccgcaaggc 840  
 atccgcacatcg gccccggcca gaccttctac gccaccgaga acatcatcg cgacatccgc 900  
 caggcccact gcaacatcg cgccggcgag tggacaagg ccgtgcagcg cgtgagcgcc 960  
 aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccagcag cggccggcgac 1020  
 ctggagatca ccacccacag ctcaactgc cgccggcgagt tcttctactg caacaccagc 1080  
 aagctgttca acagcagcta caacggcacc agtaccgcg gcaccgagag caacagcagc 1140  
 atcatcaccc tgccctgccc catcaagcag atcatcgaca tgtggcagaa ggtggccgc 1200  
 gccatctacg cccccccat cgagggcaac atcacctgca gcagcagcat caccggcctg 1260  
 ctgctggccc gcgacggcgg cctggacaac atcaccaccc agatcttccg ccccccaggc 1320  
 ggcgacatga aggacaactg ggcgcaacgag ctgtacaagt acaagggtgt ggagatcaag 1380  
 cccctggcgc tggcccccac cgagggcaag cgccgcgtgg tggagcgcga gaagcgcgcc 1440  
 gtgggcatcg ggcgcgtgat ctccggcttc ctgggcgcgg cccggcagcaa catgggcgcc 1500  
 gccagcatca ccctgaccgc ccaggccgcg cagctgctga gcggcatcg gcagcagcag 1560  
 agcaacctgc tgcgcgccat cgaggcccag cagcacatgc tgcagctgac cgtgtgggc 1620  
 atcaagcagc tgcaggcccg ctgcgtggcc atcgagcgtc acctgaagaga ccagcagctg 1680  
 ctgggcatct ggggctgcag cgccaagctg atctgcacca ccaccgtgcc ctggAACAGC 1740  
 agctggagca acaagaccca gggcgagatc tgggagaaca tgacctggat gcagtggac 1800  
 aaggagatca gcaactacac cgccatcatc taccgcctgc tggaggagag ccagaaccag 1860  
 cagggcaga acgagaagga cctgctggcc ctggacagcc gcaacaacct gtggagctgg 1920  
 ttcaacatca gcaactggct gtgttacatc aagatcttca tcatgatcg gggcggcctg 1980  
 atccgcctgc gcatcatctt cccgtgtcg agcatcgta acccgctgcg ccaggcgtac 2040  
 agccccctga gcttccagac cctgacccca aaccccccgcg gcctggaccg cctggccgc 2100  
 atcgaggagg agggcggcga gcaggaccgc gaccgcagca tccgcctggt gcaggcgttc 2160  
 ctggccctgg cctggacga ctgcgcagc ctgtgcctgt ttagctacca ccgcctgcgc 2220  
 gacctgatcc tggtgaccgc ccgcgtggtg gagctgtgg gccgcagcag ccccccggc 2280  
 ctgcagcgcg gctgggaggc cctgaagtac ctggcagcc tggtgagta ctggggcctg 2340  
 gagctgaaga agagcggcac cagcctgtg gacagcatcg ccatcgccgt ggccgaggc 2400  
 accgaccgca tcatacgaggat gatccagcgc atctaccgcg cttctgcaa catccccgc 2460  
 cgcgtgcgc agggcttcga gcccgcctg cag 2493

<210> 15  
 <211> 2565

<212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic  
 signal sequence and gp160 coding region of HIV  
 strain AF110975

<400> 15

atgcgcgtgc gcggcatcct ggcgcgtgg cagcagtggt ggatctgggg catcctggc 60  
 ttctggatct gcagcggcct gggcaacctg tgggtgaccg ttagtacgcgg cgtgcccgtg 120  
 tggcgcgagg ccagcaccac cctgttctgc gccagcgacg ccaaggccta cgagaaggag 180  
 gtgcacaacg tggggccac ccacgcctgc gtgcgcaccg accccaaacc ccaggagatc 240  
 gagctggaca acgtgaccga gaacttcaac atgtggaaa acgacatggt ggaccagatg 300  
 cacaggagaca tcatacgcc tggggaccag agcctgaagc cccgcgtgaa gctgacccca 360  
 ctgtgcgtga ccctgaagtgc caccactac agcacaact acagcaacac catgaacgac 420  
 accagctaca acaacaacac caccgaggag atcaagaact gcaccttcaa catgaccacc 480  
 gagctgcgcg acaagaagca gcaggtgtac gccctgttct acaagctggc catcggtccc 540  
 ctgaacagca acagcagcga gtaccgcctg atcaactgca acaccagcgc catcacccag 600  
 gcctggccca aggtgagctt cgaccatccat cccatccact actgcgcggcc cgccggctac 660  
 gccatcctga agtgcaagaa caacaccagc aacggcaccg gcccctgcca gaacgtgagc 720  
 accgtgcagt gcacccacgg catcaagccc gtggtgagca ccccccgtct gctgaacggc 780  
 agcctggccc agggcggcga gatcatcatc cgccagcaaga acctgagcaa caacgcctac 840

accatcatcg tgcacacctgaa cgacagcgtg gagatcgtgt gcacccgccc caacaacaac 900  
acccgcaagg gcatccgcat cggccccggc cagaccctt acgccaccca gaacatcatc 960  
ggcacatcc gccaggccca ctgcaacatc agcgccggcg agtggaaaca ggccgtgcag 1020  
cgctgagcg ccaagctgctcg ctagacttc cccaacaaga ccatcgagtt ccagcccagc 1080  
agccggcggcg acctggagat caccacccac agcttaact gccgcggcga gttcttctac 1140  
tgcaacaccca gcaagctgtt caacagcagc tacaacggca ccagctaccg cggcaccgag 1200  
agcaacagca gcatcatcac cctgcccctgc cgcataaggc agatcatcga catgtggcag 1260  
aagtgggcc gcccacatcta cggccccccc atcgagggca acatcacctg cagcagcagc 1320  
atcaccggcc tgctgtggc cggcagcggc ggcctggaca acatcaccac cgagatctc 1380  
cgcccccagg gcggcagcat gaaggacaac tggcgaacg agctgtacaa gtacaagggtg 1440  
gtggagatca agccctggg cgtggccccc accgaggcca agcgcgcgt ggtggagcgc 1500  
gagaagcgcg cctgtggcat cggcgcgtg atcttcggct tcctggcgc cggcggcagc 1560  
aacatgggcg cgcgcagcat caccctgacc gcccaggccc gccagctgct gaggcggcatc 1620  
gtcagcagc agagcaaccc gctgcgcgcc atcgaggccc agcagcacat gctgcagctg 1680  
accgtgtggg gcatcaagca gctgcaggcc cgcgtctgg ccatcgagcg ctacccatgaag 1740  
gaccagcagc tgctggcat ctgggctgc agcggcaagc tgatctgcac caccaccgtg 1800  
ccctggaaaca gcagctggag caacaagacc cagggcgaga tctgggagaa catgacctgg 1860  
atcagtggtt acaaggagat cagcaactac accggcatca tctaccgcct gctggaggag 1920  
agccagaacc agcaggagca gaacgagaag gacctgtgg ccctggacac cgcacaacaac 1980  
ctgtggagct gttcaacat cagcaacttg ctgtgttaca tcaagatctt catcatgatc 2040  
gtggcggcc tcgtccgcct ggcgcatttc ttgcgcgtc tgatcgatcgtt gatccgcgtg 2100  
cgccagggtt acagccccct gagcttccag accctgaccc ccaacccccc cggcctggac 2160  
cgctgggcc gcatcgagga ggaggcggc gagcaggacc ggcgcgcag catccgcctg 2220  
gtcagggtt tcctggccct ggctgggac gacctgcgc ggcctgtgcct gttcagctac 2280  
caccgcctgc ggcgcctgat cttgggtacc gcccgcgtgg tggagctgct gggccgcagc 2340  
agcccccgcg gcctgcagcg cggctggag gcccctgaagt acctggcag cctgggtgcag 2400  
tactggggcc tggagctgaa gaagagcgc accagcctgc tggacagcat cgcgcattgc 2460  
gtggccgagg gcaccgcaccg catcatcgag gtatccagc gcatctaccg cgccttctgc 2520  
aacatcccccc gccgcgtgcg ccaggccttc gaggccgccc tgtag 2565

<210> 16

<211> 1056

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic a gp41 coding region of HIV strain AF110975

<400> 16

gccgtggca tcggcgcgt gatcttcggc ttcctggcg cgcgcggcag caacatggc 60  
gccgcgcagca tcaccctgac cggccaggcc cgcgcgtgc tgatcgatcgt cgtgcagcag 120  
cagagcaacc tgctgcgcgc catcgaggcc cagcagcaca tgctgcagct gaccgtgtgg 180  
ggcatcaagc agctgcagcc cgcgtgtgc gccatcgagc gctacctgaa ggaccaggcag 240  
ctgtgggca tctgggctg cagcggcaag ctgatctgc ccaccaccgt gccctggAAC 300  
agcagctgga gcaacaagac ccagggcgag atctggaga acatgacctg gatgcagtg 360  
gacaaggaga tcagcaacta caccggcatc atctaccgc tgcgtggagga gagccagaac 420  
cagcaggagc agaacgagaa ggacctgtgc gcccggaca ggcgcaccaa cctgtggagc 480  
tggtaaca tcagcaactg gctgtgttac atcaagatct tcatcatgat cgtggcggc 540  
ctgatcgcc tcgcgcatttc ttgcgcgtc ctgatcgatcgt tgcgcgcgt ggcgcaggc 600  
tacagcccccc tgacttcca gaccctgacc cccaaacccc gcccggatc cccctgggc 660  
cgcatcgagg aggaggcgg cgagcaggac cgcgcaccgc gcatccgcct ggtgcaggc 720  
ttcctggccc tggcctggga cgcacgtgcg agcctgtgc tggatccatc ccaccgcctg 780  
cgcgacctga tcctgggtac cggccgcgtg gtggagctgc tggcggcag cagccccccgc 840  
ggcgcgcagc gcccggatc gcccctgaag tacctggca gcctggtgc gtaactggc 900  
ctggagctga agaagagcgc caccagcgtc ctggacagca tcgcgcattgc cgtggcggag 960

ggcaccgacc gcatcatcga ggtgatccag cgcatctacc ggcgccttctg caacatcccc 1020  
 cggcggtgc gccagggctt cgaggccgcc ctgcag 1056

<210> 17  
 <211> 492  
 <212> PRT  
 <213> Human immunodeficiency virus

<400> 17  
 Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Ala Trp  
 1 5 10 15

Glu Arg Ile Arg Leu Arg Pro Gly Gly Lys Lys Cys Tyr Met Met Lys  
 20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Lys Phe Ala Leu Asn Pro  
 35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Arg Gln Leu  
 50 55 60

His Pro Ala Leu Gln Thr Gly Ser Glu Glu Leu Lys Ser Leu Phe Asn  
 65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Glu Lys Ile Glu Val Arg Asp  
 85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Cys Gln  
 100 105 110

Gln Lys Ile Gln Gln Ala Glu Ala Ala Asp Lys Gly Lys Val Ser Gln  
 115 120 125

Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala  
 130 135 140

Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys  
 145 150 155 160

Ala Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly  
 165 170 175

Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His  
 180 185 190

Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala  
 195 200 205

Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly  
 210 215 220

Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr  
 225 230 235 240

Leu Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val  
 245 250 255

Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val  
260 265 270

Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys  
275 280 285

Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala  
290 295 300

Glu Gln Ser Thr Gln Glu Val Lys Asn Trp Met Thr Asp Thr Leu Leu  
305 310 315 320

Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly  
325 330 335

Pro Gly Ala Ser Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly  
340 345 350

Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala  
355 360 365

Asn Thr Ser Val Met Met Gln Lys Ser Asn Phe Lys Gly Pro Arg Arg  
370 375 380

Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn  
385 390 395 400

Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly  
405 410 415

His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys  
420 425 430

Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg  
435 440 445

Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr  
450 455 460

Thr Pro Gly Gln Lys Gln Glu Ser Lys Asp Arg Glu Thr Leu Thr Ser  
465 470 475 480

Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu Ser Gln  
485 490

<210> 18  
<211> 81  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic  
signal sequence of HIV strain AF110968

<400> 18

atgcgcgtga tgggcattcct gaagaactac cagcagtggg gatatgtgggg catcctggc 60  
ttctggatgc tcatcatcg c 81

<210> 19  
<211> 72  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic  
signal sequence of HIV strain AF110975

<400> 19  
atgcgcgtgc gcggcatcct gcgcagctgg cagcagtgg gatatctgggg catcctggc 60  
ttctggatct gc 72

<210> 20  
<211> 1479  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic Gag  
coding sequence of HIV strain AF110965

<400> 20  
atgggcgcgc gcccaggcat cctgcgcggc ggcaagctgg acgcctggg ggcgcattccgc 60  
ctgcgcggc gcggcaagaa gtgtacatg atgaagcacc tgggtgtggc cagccgcgag 120  
ctggagaagt tcgcctgaa ccccgccctg ctggagacca gcgagggctg caagcagatc 180  
atccgcgcgc tgacccccgc cctgcagaccc ggcagcgagg agctgaagag cctgttcaac 240  
accgtggcca ccctgtactg cgtgcacgag aagatcgagg tgcgacac caaggaggcc 300  
ctgacaaga tcgaggagga gcagaacaag tgccagcaga agatccagca ggccgaggcc 360  
gccgacaagg gcaaggtagag ccagaactac cccatctgc agaacctgca gggccagatg 420  
gtgcaccagg ccatcagccc cccgaccctg aacgcctggg tgaaggtgat cgaggagaag 480  
gccttcagcc ccgaggatgat cccatgttc accgcctgaa gcgagggccc caccggcc 540  
gacctgaaca ccatctgaa caccgtggc ggccaccagg cccatgca gatgtgaag 600  
gacaccatca acgaggaggc cccgagtg gaccgcgtgc accccgtgca cggccggcccc 660  
atcgcccccgc gccagatgcg cggccggcgc ggcagcgaca tgcggccac caccacacc 720  
ctgcaggagc agatcgccctg gatgaccagg aaccccccata tcccccgtggg cgacatctac 780  
aaggcgttga tcatcctggg cctgaacaag atcgtgcgc tgcgtacagccc cgtgacatc 840  
ctgacatca agcaggaggccc caaggaggccc ttccgcact acgtggaccg cttcttcaag 900  
accctgcgcg cccgaggcagg gtaagaact ggatgaccga caccctgctg 960  
gtcagaacgc ccaaccggc ctgcaagacc atcctgcgcg ccctggggcc cggccgcgc 1020  
ctggaggaga tgcgtaccgc ctgcgcggc gtggccggcc ccagccacaa ggccgcgtg 1080  
ctggccgagg ccatgagccca ggccaaacacc agcgtgatga tgcagaagag caacttcaag 1140  
ggcccccgcgc gcatcgaa gtcgttcaac tgcggcaagg agggccacat cggccgcac 1200  
tgcgcgcgc cccgcaagaa ggctgtgg aagtgcggca aggaggggccca ccagatgaag 1260  
gactgcaccgc agcgcaggc caacttcctg ggcaagatct ggcccagccca caaggccgc 1320  
ccggcaact tcctgcaggc cccggccggc cccaccggcc ccccccggc gagctccgc 1380  
ttcgaggaga ccaccccccgg ccagaaggcag gagacaagg accgcgagac cctgaccaggc 1440  
ctgaagagcc tgcgttcaac cggccaggcc agccagttaa 1479

<210> 21  
<211> 1509  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic Gag  
coding sequence of HIV strain AF110967

<400> 21

atgggcgccc gcgccagcat cctgcgcggc gagaagctgg acaagtggga gaagatccgc 60  
ctgcgcggcg gcggcaagaa gcactacatg ctgaagcacc tgggtgtgggc cagccgcgag 120  
ctggagggtc tcgcccgtaa ccccgccctg ctggagaccg cccgagggtctg caagcagatc 180  
atgaagcagc tgcaagccgc cctgcagacc ggcaccgagg agctgcgcag cctgtacaac 240  
accgtggcca ccctgtactg cgtgcacgac ggcacatcgagg tgcgcgacac caaggaggcc 300  
ctggacaaga tcgaggagga gcagaacaag agccagcaga agacccagca gccaaggag 360  
gccgacggca aggtgagcca gaactacccc atcgtgcaga acctgcagg ccagatggtg 420  
caccaggcca tcagcccccg caccctgaac gcctgggtga aggtgatcga ggagaaggcc 480  
ttcagccccg aggtgatccc catgttcacc gcctgagcg agggcgccac cccccaggac 540  
ctgaacacca tgctgaacac cgtggcggc caccaggccg ccatgcagat gctgaaggac 600  
accatcaacg aggaggccgc cgagtggac cgctgcacc ccgtgcaggc cggccccgtg 660  
gccccccggcc agatgcgcga ccccccggc agcgacatcg ccggcgccac cagcacccctg 720  
cagagcaga tcgcctggat gaccagcaac ccccccgtc ccgtgggca catctacaag 780  
cgctggatca tcctgggcct gaacaagatc gtgcgcatgt acagccccgt gagcatcctg 840  
gacatccgccc agggcccaa ggagccctc cgcgactacg tggaccgctt cttcaagacc 900  
ctgcgcgccc agcaggccac ccaggacgtg aagaacttgg a t g a c c g a g a c c t g 960  
cagaacgcga acccccgaactg caagaccatc ctgcgcgccc tggggcccg cgcacccctg 1020  
gaggagatga tgaccgcctg ccagggcgtg ggcggcccg gccacaaggc cgcgtgctg 1080  
gccgaggcca tgagccaggc caacagcgtg aacatcatga tgcagaagag caacttcaag 1140  
ggccccccgccc gcaacgtgaa gtgcttcaac tgcggcaagg agggccacat cgccaagaac 1200  
tgccgcgccc cccgcaagaa gggctgctgg aagtgcggca aggaggccca ccagatgaag 1260  
gactgcaccg agcgccaggc caacttcctg ggcaagatct ggcccagcca caaggccgc 1320  
cccgcaact tcctgcagaa ccgcagcgag cccgccc caccgtgcc caccgcccc 1380  
cccgccgaga gcttccgctt cgaggagacc accccccc ccaagcagga gccaaggac 1440  
cgcgagccct accgcgagcc cctgaccgccc ctgcgcagcc tgttcggcag cggccccctg 1500  
agccagtaa 1509

<210> 22

<211> 502

<212> PRT

<213> Human immunodeficiency virus

<400> 22

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Glu Lys Leu Asp Lys Trp  
1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys  
20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Gly Phe Ala Leu Asn Pro  
35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Lys Gln Ile Met Lys Gln Leu  
50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn  
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Gly Ile Glu Val Arg Asp  
85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln  
     100                        105                        110  
  
 Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn  
     115                        120                        125  
  
 Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala Ile  
     130                        135                        140  
  
 Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala  
     145                        150                        160  
  
 Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala  
     165                        170                        175  
  
 Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln  
     180                        185                        190  
  
 Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu  
     195                        200                        205  
  
 Trp Asp Arg Leu His Pro Val Gln Ala Gly Pro Val Ala Pro Gly Gln  
     210                        215                        220  
  
 Met Arg Asp Pro Arg Gly Ser Asp Ile Ala Gly Ala Thr Ser Thr Leu  
     225                        230                        240  
  
 Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Val Pro Val Gly  
     245                        250                        255  
  
 Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg  
     260                        265                        270  
  
 Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu  
     275                        280                        285  
  
 Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu  
     290                        295                        300  
  
 Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val  
     305                        310                        320  
  
 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro  
     325                        330                        335  
  
 Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly  
     340                        345                        350  
  
 Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala Asn  
     355                        360                        365  
  
 Ser Val Asn Ile Met Met Gln Lys Ser Asn Phe Lys Gly Pro Arg Arg  
     370                        375                        380  
  
 Asn Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Lys Asn  
     385                        390                        400

Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly  
405 410 415

His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys  
420 425 430

Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Asn Arg  
435 440 445

Ser Glu Pro Ala Ala Pro Thr Val Pro Thr Ala Pro Pro Ala Glu Ser  
450 455 460

Phe Arg Phe Glu Glu Thr Thr Pro Ala Pro Lys Gln Glu Pro Lys Asp  
465 470 475 480

Arg Glu Pro Tyr Arg Glu Pro Leu Thr Ala Leu Arg Ser Leu Phe Gly  
485 490 495

Ser Gly Pro Leu Ser Gln  
500

<210> 23

<211> 849

<212> PRT

<213> Human immunodeficiency virus

<400> 23

Met Arg Val Met Gly Ile Leu Lys Asn Tyr Gln Gln Trp Trp Met Trp  
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Ile Ile Ser Ser Val Val Gly Asn  
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys  
35 40 45

Thr Thr Leu Phe Cys Thr Ser Asp Ala Lys Ala Tyr Glu Thr Glu Val  
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys  
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp  
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu  
115 120 125

Lys Cys Arg Asn Val Asn Ala Thr Asn Asn Ile Asn Ser Met Ile Asp  
130 135 140

Asn Ser Asn Lys Gly Glu Met Lys Asn Cys Ser Phe Asn Val Thr Thr

145	150	155	160
Glu Leu Arg Asp Arg Lys Gln Glu Val His Ala Leu Phe Tyr Arg Leu			
165	170	175	
Asp Val Val Pro Leu Gln Gly Asn Asn Ser Asn Glu Tyr Arg Leu Ile			
180	185	190	
Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe			
195	200	205	
Asp Pro Ile Pro Ile His Tyr Cys Thr Pro Ala Gly Tyr Ala Ile Leu			
210	215	220	
Lys Cys Asn Asn Gln Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val			
225	230	235	240
Ser Ser Val Gln Cys Ala His Gly Ile Lys Pro Val Val Ser Thr Gln			
245	250	255	
Leu Leu Leu Asn Gly Ser Leu Ala Lys Gly Glu Ile Ile Ile Arg Ser			
260	265	270	
Glu Asn Leu Ala Asn Asn Ala Lys Ile Ile Ile Val Gln Leu Asn Lys			
275	280	285	
Pro Val Lys Ile Val Cys Val Arg Pro Asn Asn Asn Thr Arg Lys Ser			
290	295	300	
Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Glu Ile Ile			
305	310	315	320
Gly Asp Ile Arg Gln Ala Tyr Cys Ile Ile Asn Lys Thr Glu Trp Asn			
325	330	335	
Ser Thr Leu Gln Gly Val Ser Lys Lys Leu Glu Glu His Phe Ser Lys			
340	345	350	
Lys Ala Ile Lys Phe Glu Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr			
355	360	365	
Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asp Thr Ser			
370	375	380	
Gln Leu Phe Asn Ser Thr Tyr Ser Pro Ser Phe Asn Gly Thr Glu Asn			
385	390	395	400
Lys Leu Asn Gly Thr Ile Thr Ile Cys Arg Ile Lys Gln Ile Ile			
405	410	415	
Asn Met Trp Gln Lys Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala			
420	425	430	
Gly Asn Leu Thr Cys Glu Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg			
435	440	445	
Asp Gly Gly Lys Thr Gly Pro Asn Asp Thr Glu Ile Phe Arg Pro Gly			

450	455	460
Gly Gly Asp Met Arg Asp Asn Trp Arg Asn Glu Leu Tyr Lys Tyr Lys		
465	470	475
Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg		
485	490	495
Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe		
500	505	510
Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile		
515	520	525
Thr Leu Thr Val Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln		
530	535	540
Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln		
545	550	555
Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Ile Leu Ala Val		
565	570	575
Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser		
580	585	590
Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser		
595	600	605
Asn Arg Ser His Asp Glu Ile Trp Asp Asn Met Thr Trp Met Gln Trp		
610	615	620
Asp Arg Glu Ile Asn Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu		
625	630	635
Glu Ser Gln Asn Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu		
645	650	655
Asp Ser Trp Gln Asn Leu Trp Asn Trp Phe Ser Ile Thr Asn Trp Leu		
660	665	670
Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu		
675	680	685
Arg Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly		
690	695	700
Tyr Ser Pro Leu Pro Phe Gln Thr Leu Thr Pro Asn Pro Arg Glu Pro		
705	710	715
Asp Arg Leu Gly Arg Ile Glu Glu Gly Gly Glu Gln Asp Arg Gly		
725	730	735
Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp		
740	745	750

Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile  
755 760 765

Leu Ile Ala Ala Arg Val Leu Glu Leu Leu Gly Gln Arg Gly Trp Glu  
770 775 780

Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln Tyr Trp Gly Leu Glu Leu  
785 790 795 800

Lys Lys Ser Ala Ile Ser Leu Leu Asp Thr Ile Ala Ile Ala Val Ala  
805 810 815

Glu Gly Thr Asp Arg Ile Ile Glu Phe Ile Gln Arg Ile Cys Arg Ala  
820 825 830

Ile Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala Ala Leu  
835 840 845

Gln

<210> 24

<211> 855

<212> PRT

<213> Human immunodeficiency virus

<400> 24

Met Arg Val Arg Gly Ile Leu Arg Ser Trp Gln Gln Trp Trp Ile Trp  
1 5 10 15

Gly Ile Leu Gly Phe Trp Ile Cys Ser Gly Leu Gly Asn Leu Trp Val  
20 25 30

Thr Val Tyr Asp Gly Val Pro Val Trp Arg Glu Ala Ser Thr Thr Leu  
35 40 45

Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val His Asn Val  
50 55 60

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile  
65 70 75 80

Glu Leu Asp Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met  
85 90 95

Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu  
100 105 110

Lys Pro Arg Val Lys Leu Thr Pro Leu Cys Val Thr Leu Lys Cys Thr  
115 120 125

Asn Tyr Ser Thr Asn Tyr Ser Asn Thr Met Asn Ala Thr Ser Tyr Asn  
130 135 140

Asn Asn Thr Thr Glu Glu Ile Lys Asn Cys Thr Phe Asn Met Thr Thr

145	150	155	160												
Glu	Leu	Arg	Asp	Lys	Lys	Gln	Gln	Val	Tyr	Ala	Leu	Phe	Tyr	Lys	Leu
				165				170						175	
Asp	Ile	Val	Pro	Leu	Asn	Ser	Asn	Ser	Ser	Glu	Tyr	Arg	Leu	Ile	Asn
	180					185						190			
Cys	Asn	Thr	Ser	Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Asp
	195					200						205			
Pro	Ile	Pro	Ile	His	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys
	210					215					220				
Cys	Lys	Asn	Asn	Thr	Ser	Asn	Gly	Thr	Gly	Pro	Cys	Gln	Asn	Val	Ser
	225					230			235				240		
Thr	Val	Gln	Cys	Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Pro	Leu
				245				250					255		
Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Gly	Gly	Glu	Ile	Ile	Ile	Arg	Ser
		260					265					270			
Lys	Asn	Leu	Ser	Asn	Asn	Ala	Tyr	Thr	Ile	Ile	Val	His	Leu	Asn	Asp
		275					280					285			
Ser	Val	Glu	Ile	Val	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Gly
		290				295					300				
Ile	Arg	Ile	Gly	Pro	Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Glu	Asn	Ile	Ile
	305				310				315				320		
Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn	Ile	Ser	Ala	Gly	Glu	Trp	Asn
				325				330					335		
Lys	Ala	Val	Gln	Arg	Val	Ser	Ala	Lys	Leu	Arg	Glu	His	Phe	Pro	Asn
				340				345					350		
Lys	Thr	Ile	Glu	Phe	Gln	Pro	Ser	Ser	Gly	Gly	Asp	Leu	Glu	Ile	Thr
				355			360					365			
Thr	His	Ser	Phe	Asn	Cys	Arg	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser
				370		375					380				
Lys	Leu	Phe	Asn	Ser	Ser	Tyr	Asn	Gly	Thr	Ser	Tyr	Arg	Gly	Thr	Glu
					385		390			395			400		
Ser	Asn	Ser	Ser	Ile	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile
					405			410					415		
Asp	Met	Trp	Gln	Lys	Val	Gly	Arg	Ala	Ile	Tyr	Ala	Pro	Pro	Ile	Glu
				420			425					430			
Gly	Asn	Ile	Thr	Cys	Ser	Ser	Ile	Thr	Gly	Leu	Leu	Leu	Ala	Arg	
				435			440					445			
Asp	Gly	Gly	Leu	Asp	Asn	Ile	Thr	Thr	Glu	Ile	Phe	Arg	Pro	Gln	Gly

450	455	460
Gly Asp Met Lys Asp Asn Trp Arg Asn Glu Leu Tyr Lys Tyr Lys Val		
465	470	475
Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg		
485	490	495
Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Ile Phe		
500	505	510
Gly Phe Leu Gly Ala Ala Gly Ser Asn Met Gly Ala Ala Ser Ile Thr		
515	520	525
Leu Thr Ala Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln		
530	535	540
Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu		
545	550	555
Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Ile Glu		
565	570	575
Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly		
580	585	590
Lys Leu Ile Cys Thr Thr Val Pro Trp Asn Ser Ser Trp Ser Asn		
595	600	605
Lys Thr Gln Gly Glu Ile Trp Glu Asn Met Thr Trp Met Gln Trp Asp		
610	615	620
Lys Glu Ile Ser Asn Tyr Thr Gly Ile Ile Tyr Arg Leu Leu Glu Glu		
625	630	635
Ser Gln Asn Gln Gln Glu Gln Asn Glu Lys Asp Leu Leu Ala Leu Asp		
645	650	655
Ser Arg Asn Asn Leu Trp Ser Trp Phe Asn Ile Ser Asn Trp Leu Trp		
660	665	670
Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg		
675	680	685
Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr		
690	695	700
Ser Pro Leu Ser Phe Gln Thr Leu Thr Pro Asn Pro Arg Gly Leu Asp		
705	710	715
Arg Leu Gly Arg Ile Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg		
725	730	735
Ser Ile Arg Leu Val Gln Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu		
740	745	750

Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu  
755 760 765

Val Thr Ala Arg Val Val Glu Leu Leu Gly Arg Ser Ser Pro Arg Gly  
770 775 780

Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln  
785 790 795 800

Tyr Trp Gly Leu Glu Leu Lys Lys Ser Ala Thr Ser Leu Leu Asp Ser  
805 810 815

Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Val Ile  
820 825 830

Gln Arg Ile Tyr Arg Ala Phe Cys Asn Ile Pro Arg Arg Val Arg Gln  
835 840 845

Gly Phe Glu Ala Ala Leu Gln  
850 855

<210> 25

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 25

Asp Ile Lys Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg  
1 5 10 15

Phe Phe Lys Thr  
20

<210> 26

<211> 60

<212> DNA

<213> Human immunodeficiency virus

<400> 26

gacataaaac aaggacaaaa agagcccttt agagactatg tagaccggtt ctttaaaacc 60

<210> 27

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 27

Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg  
1 5 10 15

Phe Phe Lys Thr  
20

<210> 28  
<211> 47  
<212> PRT  
<213> Human immunodeficiency virus

<400> 28  
Thr Ile Thr Ile Thr Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln  
1 5 10 15

Lys Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Leu Thr  
20 25 30

Cys Glu Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly  
35 40 45

<210> 29  
<211> 48  
<212> PRT  
<213> Human immunodeficiency virus

<400> 29  
Ser Ile Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asp Met Trp  
1 5 10 15

Gln Lys Val Gly Arg Ala Ile Tyr Ala Pro Pro Ile Glu Gly Asn Ile  
20 25 30

Thr Cys Ser Ser Ser Ile Thr Gly Leu Leu Leu Ala Arg Asp Gly Gly  
35 40 45

<210> 30  
<211> 2469  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PR975(+)

<400> 30  
gtcgacgcca ccatggccga ggccatgagc caggccacca ggcacaacat cctgatgcag 60  
cgcaact tcaaggccc caagcgcata atcaagtct tcaactgcgg caaggaggc 120  
cacatcgccc gcaactgcgg cgccccccgc aagaagggt gctgaaatg cggcaaggag 180  
ggcaccaga tgaaggactg caccgagcgc caggccaaact tcttccgcga ggacctggcc 240  
ttcccccagg gcaaggcccq cgagttcccc agcgagcaga accgcgccaa cagccccacc 300  
agccgcgagc tgcagggtcg cgccgacaac ccccgccagcg aggccggcgc cgagcgcac 360  
ggcacccctga acttccccca gatcacccctg tggcagcgc ccctggtag catcaaggtg 420  
ggccggccaga tcaaggaggc cctgctggac accggcgccg acgacaccgt gctggaggag 480  
atgagcctgc cccgcgactg gaagccccaa atgatcgccg gcatcgccg cttcatcaag 540  
gtgcgcctgt acgaccagat cctgatcgag atctgcggca agaaggccat cggcaccgtg 600  
ctgatcgcc ccaccccccgt gaacatcatc ggccgcaca tgctgaccca gctgggctgc 660  
accctgaact tccccatcgag ccccatcgag accgtgcccgt tgaagctgaa gcccggcatg 720  
gacggcccca aggtgaagca gtggccctg accgaggaga agatcaaggc cctgaccgccc 780  
atctgcgagg agatggagaa ggaggggcaag atcaccaaga tcggcccccga gaacccttac 840  
aacaccccccgt tggtcccat caagaagaag gacagcacca agtggcgcaa gctgggtggac 900  
ttccgcgagc tgaacaagcg cacccaggac ttctggagg tgcagctggg catccccac 960

cccgcccggcc tgaagaagaa gaagagcgtg accgtgctgg acgtggcgaa cgcctacttc 1020  
 agcgtcccc tggacgagga ctccgcgaag tacaccgcct tcaccatccc cagcatcaac 1080  
 aacgagaccc ccggcatccg ctaccagtac aacgtgctgc cccaggcgtg gaagggcagc 1140  
 cccagcatct tccagagcag catgaccaag atcctggagc cttccgcgc cgcacccccc 1200  
 gagatcgtga tctaccagta catggacgac ctgtacgtgg gcagcgcacct ggagatcgac 1260  
 cagcacccgc ccaagatcga ggagctgcgc aagcacctgc tgctggggg cttcaccacc 1320  
 cccgacaaga agcaccagaa ggagcccccc ttctgttggaa tggctacgaa gctgcacccccc 1380  
 gacaagtggaa ccgtgcagcc catcgacgtc cccgagaagg agagctggac cgtgaacgac 1440  
 atccagaagc tggtggcaa gctgaactgg gccagccaga tctacccgg catcaagggtg 1500  
 cgccagctgt gcaagctgct ggcggcgcc aaggccctga cgcacatgtt gcccctgacc 1560  
 gagggggccg agctggagct ggccgagaac cgcgagatcc tgcgcgagcc cgtgcacccgc 1620  
 gtgtactacg accccagcaa ggacctggtg gccgagatcc agaaggcagg ccacgaccag 1680  
 tggacctacc agatctacca ggagcccttc aagaacctga agaccggcaa gtacgccaag 1740  
 atgcgcaccc cccacaccaa cgacgtgaag cagctgaccg aggccgtgca gaagatcgcc 1800  
 atggagagca tcgtgatctg gggcaagacc cccaagttcc gcctgcccatt ccagaaggag 1860  
 acctgggaga cctggtgac cgactactgg caggccaccc ggatccccga gtgggagttc 1920  
 gtgaacaccc ccccccgtt gaagctgtgg taccagctgg agaaggagcc catcatcgac 1980  
 gcccggaccc tctacgtggaa cggcgccgc aaccgcgaga ccaagatcg caaggccggc 2040  
 tacgtgaccg accggggccg gcagaagatc gtgagcctga ccgagaccac caaccagaag 2100  
 accgagctgc aggccatcca gctggccctg caggacagcg gcagcggaggtaaacatcg 2160  
 accgacagcc agtaccccgtt gggcatcatc caggcccgac ccgacaagag cgagacgag 2220  
 ctgtgtaacc agatcatcga gcagctgatc aagaaggaga aggtgtaccc gagctgggtg 2280  
 ccccccaca agggcatcgga cgcaacgacg cagatcgaca agctggtgag caagggcatc 2340  
 cgcaagggtgc tttccctggaa cggcatcgat ggcggcatcg tgatctacca gtacatggac 2400  
 gacctgtacg tggcagcgg cggcccttagg atcgattaaa agcttcccg ggcttagcacc 2460  
 ggtgaattc 2469

<210> 31  
 <211> 2463  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: PR975YM

<400> 31

gtcgacgcca ccatggccga ggccatgagc caggccacca ggcacccat cctgatgcag 60  
 cgcagcaact tcaaggcccc caagcgcatc atcaagtgt tcaactgcgg caaggaggc 120  
 cacatcgccc gcaactgcgg cggcccccgc aagaagggt gctggaaatg cggcaaggag 180  
 ggcaccaga tgaaggactg caccgagcgc caggccaaact tcttccgcga ggacctggcc 240  
 ttccccccagg gcaaggccccg cgagttcccc agcgagcaga accgcgcacca cagccccacc 300  
 agccgcgagc tgcaggtgcg cggcacaac cccgcagcg aggccggcgc cgagcgcac 360  
 ggcaccctga acttccccca gatcaccctg tggcagcgcc ccctggtag catcaagggt 420  
 ggcggccaga tcaaggaggc cctgctggac accggcccg acgacaccgt gctggaggag 480  
 atgagcctgc cccgcaagtg gaagcccaag atgatcgccg gcatcgccgg cttcatcaag 540  
 gtgcgcactg acgaccagat cctgatcgat atctgcggca agaaggccat cggcaccgt 600  
 ctgatcgcc cccatccccgtt gaacatcatc ggcgcacca tgctgacccca gctgggctgc 660  
 accctgaact tccccatcag cccatcgag accgtgcccc tgaagctgaa gcccggcatg 720  
 gacggccccca aggtgaagca gtggccctg accgaggaga agatcaaggc cctgaccgac 780  
 atctgcgagg agatggagaa ggagggcaag atcacaaga tcggcccccga gaaccctac 840  
 aacacccccc tttccctggaa caagaagaag gacagcacca agtggcgcaaa gctggtgac 900  
 ttccgcgagc tgaacaagcg caccgcggat ttctggggagg tgcagctggg catccccac 960  
 cccgcccggcc tgaagaagaa gaagagcgtg accgtgctgg acgtggcgaa cgcctacttc 1020  
 agcgtcccc tggacgagga ctccgcgaag tacaccgcct tcaccatccc cagcatcaac 1080  
 aacgagaccc ccggcatccg ctaccagtac aacgtgctgc cccaggcgtg gaagggcagc 1140  
 cccagcatct tccagagcag catgaccaag atcctggagc cttccgcgc cgcacccccc 1200  
 gagatcgtga tctaccagtg cccctgtac gtggcagcg acctggagat cggccagcac 1260

cgcgccaaga tcgaggagct gcgcaagcac ctgctgcgt ggggcttcac caccggac 1320  
 aagaaggcacc agaaggagcc ccccttcctg tggatgggct acgagctgca ccccgacaag 1380  
 tggaccgtgc agcccatcga gctgcccggag aaggagagct ggaccgtgaa cgacatccag 1440  
 aagctggtgg gcaagctgaa ctggggcggc cagatctacc cccggatcaa ggtgcgcccag 1500  
 ctgtgcaagc tgctgcggg cgccaaaggcc ctgaccgaca tcgtgcccct gaccgaggag 1560  
 gccgagctgg agctggccga gaaccgcggag atcctgcgcg agcccgtgca cggcgtgtac 1620  
 tacgacccca gcaaggacct ggtggccggag atccagaagc aggggccacga ccagtggacc 1680  
 taccagatct accaggagcc cttaagaac ctgaagaccg gcaagtacgc caagatgcgc 1740  
 accggccaca ccaacgacgt gaagcagctg accggaggccg tgcagaagat cgccatggag 1800  
 agcatcgta tctgggcaaa gaccccaag ttccgcctgc ccatccagaa ggagacctgg 1860  
 gagacctggt ggaccgacta ctggcaggcc acctggatcc ccgagtggga gttcgtgaac 1920  
 accccccccc tggtaagct gtggattaccag ctggagaagg agcccatcat cggcgccgag 1980  
 accttctacg tggacgcgcg cgccaaaccgc gagaccaaga tcggcaaggc cggctacgtg 2040  
 accgaccggg gccggcagaa gatcggtgac ctgaccgaga ccaccaacca gaagaccgag 2100  
 ctgcaggcca tccagctggc cctgcaggac agcggcagcg aggtgaacat cgtgaccgac 2160  
 agccagtacg ccctggcat catccaggcc cagcccgaca agagcggagag cgagctggtg 2220  
 aaccagatca tcgagcagct gatcaagaag gagaagggtgt acctgagctg ggtgcccggc 2280  
 cacaaggcga tcggcggcaaa cgagcagatc gacaagctgg tgtagcaaggg catccgcaag 2340  
 gtgctttcc tggacgcata cgatggcggc atcgatct accagtacat ggacgacctg 2400  
 tacgtggca gccccggccc taggatcgat taaaagcttc ccggggctag caccgggtgaa 2460  
 ttc 2463

<210> 32  
 <211> 2457  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: PR975YMWM

<400> 32

gtgcacgcca ccatggccga ggcacatgagc caggccacca ggcaccaacat cctgatgcag 60  
 cgcaccaact tcaaggccc caagcgcata atcaagtgt tcaactgcgg caaggaggc 120  
 cacatcgccc gcaactgccc cgccccccgc aagaagggt gctggaaatg cggcaaggag 180  
 ggcacccaga tgaaggactg caccgcgcg caggccaaact tcttcgcga ggacctggcc 240  
 ttccccccagg gcaaggcccgg cgagttcccc acgcggcaga accgcgcacaa cagcccccacc 300  
 agccgcgagc tgcagggtgcg cggcgcacaac cccgcgcg aggccggcgc cgagcgcac 360  
 ggcaccccta acttccccca gatcacccctg tggcagcgc ccctggtag catcaagggt 420  
 ggcggccaga tcaaggaggc cctgctggac accggcgcgg acgcacccgt gctggaggag 480  
 atgagcctgc cccgcgaatg gaagcccaag atgatcgccg gcatcgccgg cttcatcaag 540  
 gtgcgcctgt acgaccagat cctgatcgat atctgcggca agaaggccat cggcaccgtg 600  
 ctgatcgcc ccaacccctgt gaacatcatc ggcgcacaaca tgctgacca gctggctgc 660  
 accctgaact tccccatcag ccccatcgag accgtgcgg tgaagctgaa gccggcatg 720  
 gacggcccca aggtgaagca gtggccctgt accggaggaga agatcaaggc cctgaccgac 780  
 atctgcgagg agatggagaa ggagggcaag atcaccacaa tggccccca gAACCCCTAC 840  
 aacacccccc tggtcccat caagaagaag gacagcacca agtggcgc gctggtgac 900  
 ttccgcgagg tgaacaagcg caccaggac ttctggagg tgcagctgg catcccccac 960  
 cccgcggcc tgaagaagaa gaagagctg accgtgcgtg acgtggcga cgcctacttc 1020  
 agcgtgcggcc tggacgagga ctccgcacat tacaccgcct tcaccatccc cagcatcaac 1080  
 aacgagaccc cccgcaccc ctaccgtac aacgtgcgtc cccagggtg gaagggcagc 1140  
 cccagcatct tccagagcag catgaccaag atcctggagc cttccgcgc cgcacacccc 1200  
 gagatcgta tctaccaggc cccctgtac gtggcgcgg acctggagat cggccac 1260  
 cgcgccaaga tcgaggagct ggcacaccc ctgctgcgt ggggcttcac caccggac 1320  
 aagaaggcacc agaaggagcc ccccttcctg cccatcgacg tgcacccca caagtggacc 1380  
 gtgcagccca tcgagctgcc cgagaaggag agctggaccg tgaacgacat ccagaagctg 1440  
 gtggcgaagc tgaactggc cagccagatc taccggca tcaaggtgcg ccagctgtgc 1500  
 aagctgcgtc gccgcggccaa gcccctgacc gacatcgatc ccctgaccgaa ggaggccgag 1560

<210> 33  
<211> 9781

<212> DNA  
<213> Human immunodeficiency virus

<400> 33  
tggaaagggtt aatttactcc aagaaaaggc aagaaatcct tgatttgtgg gtctatcaca 60  
cacaaggctt ctcccgtat tggcaaaact acacaccggg gccagggtc agatatccac 120  
tgaccttgg atggtgctac aagctagtgc cagttgaccc agggaggtg gaagaggcca 180  
acggaggaga agacaactgt ttgctacacc ctatgagcca acatggagca gaggatgaag 240  
atagagaagt attaaagtgg aagtttgaca gcctccttagc acgcagacac atggcccg 300  
actacatcc ggagtattac aaagactgct gacacagaag ggactttccg cctgggactt 360  
tccactgggg cggtccggga ggtgtggct gggcgggact tggagtggt caaccctcag 420  
atgctgcata taagcagctg ctttcgcct gtactgggtc tcttcggtt gaccagatct 480  
gagcctggga gcctctggc tatctaggga acccactgct taagctcaa taaagcttc 540  
cttgagtgct ttaagtagtg tggcccatc tggtgtgtga ctctggtaac tagagatccc 600  
tcagaccctt tgggttagtg tggaaaatct ctagcagtg cgcccaaca gggaccagaa 660  
agtgaaagtg agaccagagg agatctctcg acgcaggact cggcttgctg aagtgcacac 720  
ggcaagaggc gagaggggchg gctggtgagt acgccaatt tacttgacta gcggaggcta 780  
gaaggagaga gatgggtgcg agagcgtcaa tattaagcgg cgaaaaatta gataaatggg 840  
aaagaattag gttaaaggcca gggggaaaga aacattatat gttaaaacat ctagtatggg 900  
caagcaggga gctggaaaga tttgcactta accctggcct gttagaaaca tcagaaggct 960  
gtaaacaaat aataaaacag ctacaaccag ctcttcagac aggaacagag gaacttagat 1020  
cattattcaa cacagtagca actctctatt gtgtacataa agggatagag gtacgagaca 1080  
ccaaggaagc cttagacaag atagaggaag aacaaaacaa atgtcagcaa aaagcacaac 1140  
aggcaaaagc agctgacgaa aaggtcagtc aaaattatcc tatagtacag aatgccccaaag 1200  
ggcaaatggt acaccaagct atatcaccta gaacattgaa tgcattggata aaagtaatag 1260  
agggaaaggc ttcaatcca gggaaatac ccatgtttac agcattatca gaaggagcca 1320  
ccccacaaga ttaaacaca atgttaaata cagtgggggg acatcaagca gccatgcaaa 1380  
tgtaaaaaga taccatcaat gaggaggctg cagaatggg taggacacat ccagtacatg 1440  
cagggcctgt tgaccaggc cagatgagag aaccaagggg aagtgcacata gcaggaacta 1500  
ctagtaccct tcaggaacaa atagcatgga tgacaagtaa tccacctatt ccagtagaaag 1560  
acatctataa aagatggata attctgggt taaataaaat agtaagaatg tatagccctg 1620  
ttagcatttt ggacataaaa caaggccaa aagaaccctt tagagactat gtagaccgg 1680  
tcttaaaaac cttaaagagct gaacaagcta cacaagatgt aaagaattgg atgacagaca 1740  
ccttgttggc cccaaatgct aacccagatt gtaagaccat ttaagagca ttaggaccag 1800  
ggcctcatt agaagaaatg atgacagcat gtcagggagt gggaggacct agccataaaag 1860  
caagagtgtt ggctgaggca atgagccaag caaacagtaa catactagtg cagagaagca 1920  
atttaaagg ctctaacaga attattaaat gttcaactg tggcaaagta gggcacatag 1980  
ccagaaattg cagggccctt agggaaaagg gctgttgaa atgtggacag gaaggacacc 2040  
aatgaaaga ctgtactgag aggccaggcta attttttagg gaaaatttqg cttcccaaca 2100

agggaggccc agggaaatttc ctccagaaca gaccagagcc aacagccccca ccagcagaac 216  
caacagcccc accagcagag agcttcagg tgcaggagac aaccccccgtg ccgaggaagg 2220  
agaaaagagag ggaaccttta acttcctca aatcactttt tggcagcgac cccttgctc 2280  
aataaaagta gagggccaga taaaggaggc tctcttagac acaggagcag atgatacagt 2340  
attagaagaa atagatttgc cagggaaatg gaaaccaaaa atgatagggg gaattggagg 2400  
ttttatcaa gtaagacagt atgatcaa atctatagaa atttggaa aaaaggctat 2460  
aggtacagta tttagtagggc ctacaccagt caacataattt ggaagaaatc tgtaactca 2520  
gcttggatgc acactaaattt ttccaatttgc tccttattgaa actgttaccag taaaattaaa 2580  
accaggaatg gatggcccaa aggtcaaaca atggccattt acagaagaaa aaataaaagc 2640  
attaacagca atttgtgagg aaatggagaa ggaaggaaaa attacaaaaaa ttgggcctga 2700  
taatccatat aacactccag tatggccat aaaaaaagaag gacagtacta agtggagaaa 2760  
atttagtagat tttagggaaac tcaataaaag aactcaagac ttttggaaat ttcaattttagg 2820  
aataccacac ccagcaggat taaaagaa aaaatcagtg acagtgcgt atgtggggga 2880  
tgcataat ttcagttcctt tagatgaaag cttcaggaaa tatactgc taccatacc 2940  
tagtataaac aatgaaacac cagggattt atatcaat taatgtctgc cacaggatg 3000  
gaaaggatca ccagcaatata tccagatg catgacaaaa atcttagagc cttcagagc 3060  
aaaaaaatcca gacatagttt tctatcaata tatggatgac ttgtatgtat gatctgactt 3120  
agaaaataggg caacatagag caaaaataga agagtttggg gaacatttt tgaaatgggg 3180  
atttacaaca ccagacaaga aacatcaaaa agaacccca ttttttggg tgggtatga 3240  
actccatcct gacaaatggg cagtacaacc tatactgc tccagaaaagg atagttggac 3300  
tgtcaatgtat atacagaagt tagtggaaa attaaactgg gcaagtca gttaccagg 3360  
gattaaagta aggcaactct gttaactcct cagggggggc aaagcactaa cagacatagt 3420  
accactaact gaagaagcag aatttgcattt ggcagagaac agggaaattt taagagaacc 3480  
agtacatgga gtatattatg atccatcaaaa agacttgata gctgaaatac agaaacaggg 3540  
gcatgaacaa tggacatatac aaatttatca agaaccattt aaaaatctga aaacaggaa 3600  
gtatgcaaaa atgaggacta cccacactaa tgatgtaaaa cagttacag aggcagtgc 3660  
aaaaatagcc atggaaagca tagtaatatg gggaaagact cctaaattt gactaccat 3720  
ccaaaaagaa acatgggaga catggggac agactattgg caagccaccc ggtatccctga 3780  
gtgggagttt gttaataccctt ctcccttagt aaaattatgg taccactag aaaaagatcc 3840  
catagcagga gttagaaactt tctatgtata tggagcaact aataggaaatg ctaaaatagg 3900  
aaaagcaggg tatgttactg acagaggaag gcagaaaattt gttactctaa ctaacacaac 3960  
aaatcagaag actgagttac aagcaattca gtagctctg caggattcag gatcagaagt 4020  
aaacatagta acagactcac agtatgcattt aggaatcattt caagcacaac cagataagag 4080  
tgactcagag atatttaacc aaataataga acagttata aacaaggaaa gatatctacct 4140  
gtcatggta ccagcacata aaggaattgg gggaaatgaa caagtagata aatttagtaag 4200  
taagggattt agaaaagtgt tgttctaga tggaaatagat aaagctcaag aagagcatga 4260  
aaggtaccac agcaattggg gagcaatggc taatgagttt aatgtccac ccatagtagc 4320  
aaaagaaata gtagctagct gtgataatg tcaatggaaa gggaaagcca tacatggaca 4380  
agtcgactgt agtccaggaa tatggcaattt agattgtacc catttagagg gaaaatcat 4440  
cctggtagca gtccatgtat ctatggcta catggaaagca gaggttatcc cagcagaaac 4500  
aggacaagaa acagcatattt ttatattaaa attagcagga agatggccag tcaaagtaat 4560  
acatacagac aatggcagta attttaccag tactgcagg aaggcagcc gttggtggc 4620  
aggtatccaa caggaatttg gaattccctt caatccccaa agtcagggag tggtagaaatc 4680  
catgaataaa gaattaaaga aaataatagg acaagtaaga gatcaagctg agcaccttaa 4740  
gacagcagta caaatggcag tattcattca caatttaaa agaaaagggg gaattggggg 4800  
gtacagtgc gggaaagaa taatagacat aatagcaaca gacataaaaaa ctaaaagaatt 4860  
acaaaaaacaa attataagaa ttcaaaat tccgggtttt tacagagaca gcagagaccc 4920  
tatttggaaa ggaccagccc aactactctg gaaaggtgaa ggggttagtag taatagaaga 4980  
taaaggtgac ataaaaggtag tccaaggag gaaagcaaaa atcatttagag attatggaaa 5040  
acagatggca ggtgctgatt gtgtggcagg tggacaggat gaagattaga gcatggaaata 5100  
gtttagtaaa gcaccatattt tataatcaa ggagagctag tggatggtc tacagacatc 5160  
attttggaaag cagacatcca aaagtaagtt cagaagtaca tatcccattt gggatgcta 5220  
gatttagtaat aaaaacatattt tgggtttgc agacaggaga aagagattgg catttggc 5280  
atggagtc catagaatgg agactgagag aatacagcac acaagtagac cctgacccctt 5340  
cagaccagct aattcacatcg cattattttgc attgttttac agaatctgcc ataagacaag 5400  
ccatatttagg acacatagttt tttccttagt gtgactatca agcaggacat aagaaggtag 5460  
gatctctgca atacttggca ctgacagcat tgataaaaacc aaaaaaqaga aqccaccc 5520

tgccatgt tagaaaatta gtagaggata gatggAACGA cccccAGAAg accAGGGGCC 5580  
 gcagAGGGAA ccatacaatg aatggacact agagattcta gaagaACTCA aGcAGGAAGC 5640  
 tgcAGACAC ttccTAGAC catggCTCA tagCTTAGGA caatATATCT atgAAACCTA 5700  
 tggggataCT tggACGGGAG ttGAAGCTAT aataAGAGTA ctGCAACAC tactGTTCAT 5760  
 tcatttcAGA attggatGCC aacatAGCAG aatAGCAtC ttGCGACAGA gaAGAGCAAG 5820  
 aaatggAGCC agtagatCCT aaactAAAGC cctggAACCA tccAGGAAGC caACCTAAA 5880  
 cAGCTTGTAA taattGCTT tgcaAAACACT gtAGCTATCA ttGTCAGTT tgCTTCAGA 5940  
 caaaAGGTTT aggCATTCC tatGGCAGGA agaAGCGGAG acAGCgACGA aGcGCTCCTC 6000  
 caAGTGGTGA agatcatCAA aatCCTCTAT caaAGCAGTA agtACACATA gtagATGTA 6060  
 tggtaAGTTT aagtTTATTt aaaggAGTAG attATAGATT aggAGTAGGA gcATTGATAG 6120  
 tagcACTAAT catAGCAATA atAGTGTGGA ccatAGCATA tatAGAATAT aggAAATTGG 6180  
 taagacAAAAA gaaaATAGAC tggTTAATTa aaAGAAATTAG ggAAAGAGCA gaAGACAGTG 6240  
 gcaatgAGAG tgatGGGAC acAGAAGAAt tgtcaACAAt ggtGGATATg gggcatCTTA 6300  
 ggcttCTGGA tgctaATGAT ttGTAACACG gaggACTGT gggTCACAGT ctACTATGGG 6360  
 gtacCTGTGT ggAGAGAAgC AAAAActACT ctATTCTGT catCAGATGC taaAGCATAT 6420  
 gagacAGAGAG tgcATAATGT ctGGGCTACA catGCTGTG tacCCACAGA ccccaACCCa 6480  
 caagAAATAG ttttGGGAAA tGtaACAGAA aatttTAATA tGtGAAAaaa taACATGGCA 6540  
 gatcAGATGc atGAGGATAT aatcAGTTA tGGATCAA GcCTAAAGCC atGtGTAAG 6600  
 ttGACCCAC tctGTGTCAC ttAAACTGT acAGATAAA atGTTACAGG taATAGAACT 6660  
 gttacAGGTA atacaAAATGA tACCAATATT gcaaATGCTA catATAAGTA tgaAGAAATG 6720  
 AAAAAttGCT ctTTCAATGC aaccACAGAA ttaAGAGATA agAAACATAA agAGTATGCA 6780  
 ctCTTTATA aactTGATAT agTACCACTT aatGAAAATA gtaACAACTT tacATATAGA 6840  
 ttaataAAATT gcaATACCTC aaccATAACA caAGCCTGTC caaAGGTCTC tttGACCCG 6900  
 attcCTATAc attACTGTGc tccAGCTGAT tatGCGATTc taaAGTGTAA taATAAGACA 6960  
 ttcaATGGGA caggACCATG ttATAATGTC agcACAGTAC aatGtACACA tGGAATTAAg 7020  
 ccAGTGGTAT caACTCAACT actGTTAAAT ggtAGTCTAG cAGAAGAAGG gataATAATT 7080  
 agatCTGAAA attTGACAGA gaATAACAAA acaATAATAG tacATCTAA tGAATCTGTA 7140  
 gagattaATT gtacaAGGCC caACAATAAT acaAGGAAAAA gTGTAAAGGAT aggACCAGGA 7200  
 caAGCATTCT atGCAACAAA tgACGTAATA ggAAACATAA gacaAGCACA ttGTAACATT 7260  
 agtACAGATA gatGGAAATAA aactTTACAA cAGGTAATGA AAAAAttAGG agAGCATTc 7320  
 CCTAAATAAA CAATAAAATT tGAACCACAT GcAGGAGGG AtCTAGAAAT tacaATGcAT 7380  
 agCTTAATT gtagAGGAGA ATTtTCTAT tgcaATACAT caAAACCTGT taATAGTACA 7440  
 tactACCTA agaATGGTAC atACAATAAC aatGGTAATT caAGCTTACc catCACACTC 7500  
 caATGCAAAA taaaACAAAT tGtACGCTAG tggCAAGGG tagGACAAGC aatGtATGCC 7560  
 CCTCCCATTG caggAAACAT aacATGTAGA tcaaACATCA cAGGAATACT attGACACGT 7620  
 gatGGGGAT ttaACAACAC aaACAACGAC acAGAGGAGA cATTcAGACC tggAGGAGGA 7680  
 gatATGAGGg ataACTGGAG aagtGAATTt tataAAATATA aagtGGTGA aAtTAAGCCA 7740  
 ttGGAATAG cacCCACTA gGCAAAAGA agAGTGGTGC agAGAAAAAA aAGAGCAGTG 7800  
 ggaATAGGAG ctGTGTTcCT tggTTCTG ggAGCAGCAG gaAGCACTAT gggcgcAGCG 7860  
 tcaataACGC tgacGGTACA gGCCAGACAA ctGTTGTCtG gtATAGTGCa acAGCAAAGC 7920  
 aatttGCTGA aggCTATAGA ggcGCAACAG catATGTGc aactCACAGT ctGGGGCATT 7980  
 aAGCAGCTCC aggCGAGAGT CCTGGCTATA gaaAGATAcc taaAGGATCA acAGCTCCTA 8040  
 gggATTTGGG gctGCTCTGG aagACTCATC tgcACCACTG ctGTGCTTg gaACTCCAGT 8100  
 tggAGTAATA aatCTGAAGC agatATTGG gataACATGA ctTGGATGCA gtGGGATAGA 8160  
 gaaATTAATA attACACAGA aacaATATTc agGTTGCTTg aAGACTCGCA aaACCAGCAG 8220  
 gaaaAGAATG AAAAAGATTt attAGAATTG gacaAGTGGa ataATCTGTG gaATTGGTT 8280  
 gacATATCAA actGGCTGTG gtATATAAAA atATTcATAA tgATAGTAGG aggCTTgATA 8340  
 ggttAAAGAA taattttGc tGtGCTCTCt atAGTGAATA gAGTGGAGCA gggATACTCA 8400  
 CCTTTGTcAT ttcaGACCCt tACCCCAAGC ccGAGGGGAG tcGACAGGCT cggAGGAATC 8460  
 gaAGAAGAAG gtGGAGAGCA agACAGAGAC agATCCATAC gATtGGTGAg cggATTCTG 8520  
 tcGCTTGCCT gggACGATCT gcccAGCCTG tgcCTTCA gCTAccACCG ctTGAGAGAC 8580  
 ttCATATTAA ttGcAGTGTAG gGcAGTGGAA ctTCTGGAC acAGCAGTCT cAGGGGACTA 8640  
 cAGAGGGGT gggAGATCCT taAGTATCTG gGAAGTCTG tGcAGTATTG gggTCTAGAG 8700  
 ctaaaaaAGA gTGCTATTAG tccGCTTGTAT accATAGCAa tagCAGTAGC tGAAGGAACA 8760  
 gatAGGATTA tagAAATTGt acaaAGAAATT tGtAGAGCTA tcCTCAACAT acCTAGGAGA 8820  
 ataAGACAGG gCTTTGAAGC agCTTGTcTA taaaATGGGA ggCAAGTGGT caAAACGcAG 8880  
 catAGTGGA tggCCTGcAG taAGAGAAAG aatGAGAAGA actGAGCCAG cAGCAGAGGG 8940

agtaggagca gcgtctcaag acttagatag acatgggc a ttacaagca gcaacacacc 9000  
 tgctactaat gaagcttgc cctggctgca agcacaagag gaggacggag atgtaggctt 9060  
 tccagtca g ctcaggta cttaaagacc aatgacttat aagagtgcag tagatctca 9120  
 cttctttta aaagaaaagg gggactgga agggttaatt tactcttagga aaaggcaaga 9180  
 aatccttgc ttgtgggtct ataacacaca aggcttctc cctgattggc aaaactacac 9240  
 atcggggcca ggggtccgat tcccactgac ctggatgg tgctcaagc tagtaccagt 9300  
 tgacccaagg gaggtgaaag aggccatga aggagaagac aactgtttgc tacaccstat 9360  
 gagccacat ggagcagagg atgaagatag agaagtatta aagtggaaat ttgacagcct 9420  
 tctagcacac agacacatgg cccgcgagct acatccggag tattacaaag actgctgaca 9480  
 cagaaggac ttccgcctg ggacttcca ctggcggtt ccgggagggtg tggctggc 9540  
 gggacttggg agtggtcacc ctcagatgct gcatataagc agtgccttt cgctgtact 9600  
 gggctctct cggtagacca gatctgagcc tggagctct ctggctatct agggAACCCA 9660  
 ctgcttaggc ctcataaaag cttgccttga gtgctctaag tagtgtgtgc ccattgttg 9720  
 tgtgactctg gtaactagag atccctcaga cccttgc tggatgtggaa aatctctagc 9780  
 a 9781

<210> 34  
 <211> 203  
 <212> DNA  
 <213> Human immunodeficiency virus

<400> 34  
 gctgaggcaa tgagccaagc aaccagcgca aacatactga tgcagagaag caatttcaaa 60  
 ggcctaaaaa gaattattaa atgtttcaac tgtggcaagg aagggcacat agctagaaat 120  
 tgttagggccc ctagaaaaaa aggctgttgg aatgtggaa aggaaggaca ccaaattgaaa 180  
 gactgtactg agaggcaggc taa 203

<210> 35  
 <211> 2151  
 <212> DNA  
 <213> Human immunodeficiency virus

<400> 35  
 ttttttaggg aagatttggc ctccccacaa gggaggc a gggatttcc ttca gaaacag 60  
 aacagagcca acagccccac cagcagagag cttcaagttc gaggagacaa cccccgctcc 120  
 gaagcaggag ccgaaagaca gggaaaccctt aattttccctc aaatcactct ttggcagcga 180  
 ccccttgc tcaataaaatg agggtgtcaa ataaaggagg ctctcttgc cacaggagct 240  
 gatgatacag tattagaaga aatgagtttgc ccaggaaaat gggaaacccaa aatgatagga 300  
 ggaatttggag gttttatcaa agtaagacag tatgatcaaa tacttata gaaatgttgg 360  
 aaaaaggcta taggtacagt attaataatg cttcacacctg tcaacataat tggaaaggaat 420  
 atgttgactc agcttggatg cacactaaat tttccaaat gttccatttgc aactgtgcc 480  
 gtaaaattaa agccagaat ggtggccca aaggttaaac aatggccatt gacagaagag 540  
 aaaataaaatg cattaacacg aattttgtgaa gaaatggaga aagaaggaaa aattacaaa 600  
 attgggcctg aaaatccata taacacttca gtatttgccca taaaaaaatgaa ggacagtact 660  
 aagtggagaa agtttagtata tttcaggaa cttataaaaaa gaaatcaaga cttttggaa 720  
 gttcaatttgc gaataaccaca cccagcagg tttaaaaatgaa aaaaatcagt gacagtactg 780  
 gatgtggggg atgcataattt ttcagttcct ttagatgagg acttcaggaa atataactgca 840  
 ttccacccata ctagtataaa caatgaaaca ccaggatata gatataatcaata taatgtgc 900  
 ccacaggat gggaaaggatc accatcaata ttccagatgat gcatgacaaa aatcttagag 960  
 ccctttagag caagaaatcc agaaatagtc atctatcaat atatggatgat gttgtatgt 1020  
 ggtctgact tagaaatagg gcaacataga gcaaaaaatag aggatgttggaaaatctg 1080  
 ttaaggtggg gatttaccac accggacaag aaacatcaga aagaacccccc atttttttgg 1140  
 atggggatg aactccatcc tgacaaatgg acatgacgc ctatagatgtt gccagaaaag 1200  
 gaaagctgga ctgtcaatga tatacagaag ttatgtggaa aattnaaatttggggcc 1260  
 atttacccatg gattnaaatg aaggcaactt tgtaaaactcc tttagggggcc caaagcacta 1320  
 acagatatacg taccactaac tgaagaagca gattnaaatg tggcagagaa caggaaattt 1380  
 ctaagagaac cagtagatgg agtataattt gacccatcaa aagacttggt agctgaaata 1440

cagaaacagg ggcatgacca atggacatat caaatttacc aagaaccatt caaaaacctg 1500  
aaaacaggga agtatgc aaaatgaggact gcccacacta atgatgtaaa acagttaca 1560  
gaggcagtgc aaaaaatacg tatggaaagc atagtaatat ggggaaagac tcctaaattt 1620  
agactaccca tccaaaaaga aacatgggag acatggtggc cagactattt gcaagccacc 1680  
tggattcctg agtgggagtt ttataacc cctcccttag taaaattatg gtaccagcta 1740  
gagaaagaac ccataatagg agcagaaact ttctatgttag atggagcagc taataggaa 1800  
actaaaatag gaaaaggcagg gtatgttact gacagaggaa ggcagaaaat tgtttctcta 1860  
acagaaacaa caaatcagaa gactgaatta caagcaattc agctagctt gcaagattca 1920  
ggatcagaag taaacatagt aacagactca cagtatgcat taggaatcat tcaagcacaa 1980  
ccagataaga gtgaatcaga gtttgtcaac caaataatag aacaattaat aaaaaaggaa 2040  
aaggtctacc tgtcatgggt accagcacat aaaggaattt gaggaaatga acaaataatag 2100  
aaattagtaa gtaagggaa caggaaagtg ctgtttcttag atgaaataga t 2151

<210> 36  
<211> 54  
<212> DNA  
<213> Human immunodeficiency virus

<400> 36  
ggccgcatcg tgcgttacca gtacatggac gacctgtacg tggcagcgg cgcc 54

<210> 37  
<211> 18  
<212> PRT  
<213> Human immunodeficiency virus

<400> 37  
Gly Gly Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser  
1 5 10 15

Gly Gly

<210> 38  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer  
S1FCSactA

<400> 38  
gtttcttgag ctctggaagg gttaatttac tccaaagaa 38

<210> 39  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer  
S1FTSacTA

<400> 39  
gtttcttgag ctctggaagg gttaatttac tctaagaa 38

```

<210> 40
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
      S145RTSalTA

<400> 40
gtttcttgtc gacttgtcca tgtatggctt cccct 35

<210> 41
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
      S145RCSalTA

<400> 41
gtttcttgtc gacttgtcca tgcattggctt ccct 34

<210> 42
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
      S245FASalTA

<400> 42
gtttcttgtc gactgttagtc caggaatatg gcaattag 38

<210> 43
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
      S245FGSalTA

<400> 43
gtttcttgtc gactgttagtc caggatatg gcaattag 38

<210> 44
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
      S2FullNotTA

```

<400> 44  
gttcttgcg gccgctgcta gagattttcc acactacca

39

<210> 45  
<211> 9738  
<212> DNA  
<213> Human immunodeficiency virus

<400> 45  
tggaaagggtt aatttactcc aggaaaaggc aagagatcct tgatttatgg gtctatcaca 60  
cacaaggcta cttccctgat tggcaaaact acacaccggg accaggggtc agatatccac 120  
tgacctttgg atgggtgttc aagctagtgc cagttgaccc aagggaagta gaagaggcca 180  
acgaggaga agacaactgt ttgctacacc ctatgagcca gtatggaatg gatgatgaac 240  
acaagaagaat gttacagtgg aagtttgaca gcagccttagc acgcagacac ctggcccg 300  
agctacatcc ggattattac aaagactgct gacacagaag ggacttccg cctgggactt 360  
tccactgggg cggttccaggg ggagtggct gggcgggact gggagtggcc agccctcaga 420  
tgctgcataat aagcagcggc tttcgcctg tactgggtct ctctaggtag accagatccg 480  
agctctggag ctctctgtct atctggggaa cccactgctt aggccctcaat aaagcttgcc 540  
ttgagtgtct taagtagtgt gtgcccatct gttgtgtac tctggtaact ctggtaacta 600  
gagatccctc agaccctttg tggtagtgt gaaaatctct agcagtggcg cccgaacagg 660  
gacttggaaag cgaaaatgtg accagagaag atctctcgac gcaggactcg gcttgctgaa 720  
gtgcactcgg caagaggcga gggggcgac tggtagtac gccaaaattt tttttgacta 780  
gcggaggcta gaaggagaga gatgggtgcg agagcgtcaa tattaagagg gggaaaatta 840  
gacaaatggg aaaaaattag gttacggcca gggggagaa aacactataat gctaaaacac 900  
ctagtatggg caagcagaga gctggaaaga tttgcagttt accctggcct ttttagagaca 960  
tcagacggat gtagacaat aataaaacag ctacaaccag ctcttcagac aggaacagag 1020  
gaaatttagat cattatttaa cacagtagca actctctatt gtgtacataa agggatagat 1080  
gtacgagaca ccaaggaagc cttagacaag atagaggagg aacaaaacaa atgtcagcaa 1140  
aaaacacacgc aggcggaagc ggctgacaaa aaggtcagtc aaaattatcc tatagtgcag 1200  
aacctccaag ggcaaatggt acaccaggcc atatcaccta gaaccttggaa tgcattggta 1260  
aaagtaatag aggagaaggc ttttagccca gaggtataac ccatgtttac agcattatca 1320  
gaaggagcca cccccacaaga tttaaacacc atgttaataa cagtgggggg acatcaagca 1380  
gccatgcaaa tggtaaaaaga taccatcaat gaggaggctg cagaatggga tagttacat 1440  
ccagtacatg cagggcctgt tgaccaggc cagatgagag aaccaagggg aagtgacata 1500  
gcaggaacta ctagtaccct tcaagaacaa atagcatgga tgacaagtaa cccacctatc 1560  
ccagtagggg acatctataa aaggtggata attctgggt taaataaaat agtaagaatg 1620  
tacagccctg tcagcatttt agacataaaa caaggacca aggaaccctt tagagactat 1680  
gtagaccggc tcttcaaaaac tttaagagct gaacaatcta cacaagaggt aaaaaattgg 1740  
atgacagaca ccttgttagt cccaaatgcg aacccagatt gtaagaccat tttaagagca 1800  
ttaggaccag gggcttcatt agaagaaatg atgacagcat gtcaggggagt gggaggacct 1860  
agccacaaag caagagttt ggctgaggca atgagccaag caaacaatac aagtgtaatg 1920  
atacagaaaaa gcaattttaa aggcctaga agagctgtt aatgtttcaa ctgtggcagg 1980  
gaaggcaca tagccaggaa ttgcaggggc cctaggaaaa ggggctgtt gaaatgtgga 2040  
aaggaaaggac accaaatgaa agactgtact gagaggcagg ctaattttt agggaaaatt 2100  
tggccttccc acaagggag gccaggaaat ttccctcaga gcagaccaga gccaaacagcc 2160  
ccaccactag aaccaacagc cccaccagca gagagctca agttcaaggaa gactccgaag 2220  
cagagccga aagacaggga acctttaact tccctcaat cactcttgg cagcgacccc 2280  
ttgtctcaat aaaagttagcg ggc当地acaa aggaggctt ttagataca ggagcagatg 2340  
atacagtact agaagaaata aacttgcag gaaaatggaa accaaaaatg ataggaggaa 2400  
tttgagggtt tatcaaagta agacagtatg atcaaataact tatagaatatt tggaaaaaaa 2460  
ggctatagg tacagtatta gtaggaccta cacctgtcaa cataattggg agaaatctgt 2520  
tgactcagct tggatgcaca ctaaattttc caatttagccc cattgaaact gtaccagtaa 2580  
aattaaagcc aggaatggat ggcccaaagg ttaaacaatg gccattgaca gaagaaaaaaa 2640  
taaaagcatt aacagaaaatt tggagggaa tggagaagga agggaaaatt aaaaaattg 2700  
ggcctgaaaaa tccatataac actccagttt ttgcataaa gaagaaggac agtacaaatg 2760

ggagaaaatt agtagatttc aggaaactca ataaaagaac tcaagacttt tggaaagtcc 2820  
 aatttagaat accacaccca gcagggttaa aaaagaaaaaa atcagtgaca gtactggatg 2880  
 tggagatgc atattttca gtcctttag atgagagctt cagaaaatat actgcattca 2940  
 ccatacctag tataaacaat gaaacaccag ggattagata tcaatataat gttctccac 3000  
 aggatggaa aggatcacca gcaatattcc agagtagcat gacaagaatc ttagagccct 3060  
 ttagaacaca aaacccagaa gtagttatct atcaatatat ggatgactt tatgtaggat 3120  
 ctgacttaga aataggcaa catagagcaa aaatagagga gttaaagagga cacctattga 3180  
 aatggggatt taccacacca gacaagaaac atcagaaaaga acccccattt ctttggatgg 3240  
 ggtatgaact ccatcctgac aaatggacag tacagcctat acagctgccaa gaaaaggaga 3300  
 gctggactgt caatgatata cagaagttag tggaaagtt aaactggca agtcagattt 3360  
 acccaggat taaagtaagg caactgtgta aactccttag gggagccaa gcactaacag 3420  
 acatagtgcc actgactgaa gaagcagaat tagaattggc tgagaacagg gaaattctaa 3480  
 aagaaccagt acatggagta tattatgacc catcaaaaaga ttaatagct gaaatacaga 3540  
 aacagggaa tgaccaatgg acatataaaa ttaccaaga accattaaa aatctgagaa 3600  
 cagaaagta tgcaaaaatg aggactgccc acactaatga tgtgaaacag ttagcagagg 3660  
 cagtcaaaa gataaccagg gaaagcatag taatatgggg aaaaactcct aaatttagac 3720  
 taccatccc aaaagaaaaca tggagacat ggtggcaga ctattggcaa gccacctgga 3780  
 ttcttgatg gtagttgtc aataccctc ccctagtaaa atttgtgtac cagctggaaa 3840  
 aagaacccat agtagggca gaaacttct atgtagatgg agcagccat agggaaacta 3900  
 aaataggaaa agcagggtat gtcactgaca aaggaaggca gaaagttgt tccttactg 3960  
 aaacaacaaa tcagaagact gaattacaag caatttagt agctttgcag gattcaggc 4020  
 cagaagtaaa catagtaaca gactcacagt atgcattagg aatcattcaa gcacaaccag 4080  
 ataagagtga atcagaatta gtcagtcaaa taatagaaca gttgataaaa aaggaaaaag 4140  
 tctacctatc atgggtacca gcacataaaag gaattggagg aaatgaacaa gtagacaaat 4200  
 tagtaagtag tggaaatcaga aaagtactgt ttcttagatgg aatagataaa gctcaagaag 4260  
 agcatgaaaa atatcacagc aattggagag caatggctag tgagtttaat ctgccaccca 4320  
 tagtagcaaa ggaaatagta gccagctgtg ataaatgtca gctaaaaggga gaagccatgc 4380  
 atgacaagt cgactgtat ccaggaatat ggcaattaga ctgtacacat tttagaaggaa 4440  
 aaatcatcct agtagcagtc catgtgcca gtggctacat ggaaggcagag gttatcccag 4500  
 cagaaacagg acaagaaaca gcatactta tactaaaatt agcaggaaga tggccagtca 4560  
 aagtaataca tacagataat ggtagtaatt tcaccagtac cgcagttaaag gcagcctgtt 4620  
 ggtggcaga tatccaacgg gaatttgaa ttccctacaa tccccaaagt caaggagtag 4680  
 tagaatccat gaataaaagaa taaaagaaaa tcataggca agtaagagat caagctgagc 4740  
 acctaagac agcagtacaa atggcgtat tcattcacaat tttaaaaga aaagggggg 4800  
 ttgggggta cagtgcaggg gagagaataa tagacataat agcatcagac atacaaacta 4860  
 aagaattaca aaaacaaatt ataaaaatttca aaaattttcg gtttattac agagacagca 4920  
 gagaccctat ttggaaagga ccagccaaac tactctggaa aggtgaaggg gcagtagtaa 4980  
 tacaagataa tagtgatata aaggttagtac caagaaggaa agcaaaaatc attaaggact 5040  
 atgaaaaca gatggcaggt gctgattgtg tggcaggtag acaggatgaa gattagaaca 5100  
 tggcacagtt tagtaaagca ccatatgtat gtttcgagga gagctgatgg atggttctac 5160  
 agacatcatt atgaaagcag acacccaaaa gtaagttcg aagtacacat cccattagga 5220  
 gatgccaggt tagtaataaa aacatattgg ggtctgcaga caggagaaag agcttggcat 5280  
 ttgggtcacf gtagtccat agaatggaga ttgagaagat atagcacaca agtagaccct 5340  
 gacctgacag accaactaat tcatacgat tattttgatt gtttgcaga atctgccata 5400  
 agggaaagcca tactaggaca gatagttgc cctaagtgtg actatcaagc aggacataac 5460  
 aaggttagat ctctacaata ctggcactg acagcattga taaaaccaaa aaagataaag 5520  
 ccacctctgc ctagttagt gaaatttagta gaggatagat ggaacaagcc ccagaagacc 5580  
 agggccgca gagggaaacca tacaatgaat ggacactaga gcttttagaa gaactcaagc 5640  
 aggaagctgt cagacacttt cctagaccat ggctccatata cttaggacaa catatctatg 5700  
 aaacctatgg agatacttgg acaggagttt aagcaataat aagaatcctg caacaattac 5760  
 tgtttattca ttccagatt ggtgcacat atagcagaat aggcatggc cgacagagaa 5820  
 gagcaagaaa tggagccat agatcctaacc ctagaaccct ggaaccatcc aggaagttag 5880  
 cctaaaactg cttgtatgg gtgttactgt aaacgttgca gctatcattt tctagttgc 5940  
 ttccagaaaa aaggcttagg catttactat ggcaggaaga agcggagaca gcgacgaagc 6000  
 gctccctccaa gcaataaaaga tcatcaagat cctctaccaa agcagtaagt accgaatagt 6060  
 atatgtaatg ttagattaa ctgcaagaat agattctaga ttaggaatag gacgattgtat 6120  
 agtagcacta atcatagcaa taatagtgtg gaccatagta tatatagaat atagggaaatt 6180

ggttaaggcaa aggaaaatag actggtagt taaaaggatt agggaaagag cagaagacag 6240  
 tggcaatgag agcgaggggg atactgaaga attatcgaca ctggtgata tggggcatct 6300  
 tagctttt gatgctaatg atgtgtatg tgaaggcctt gtgggtcaca gtctactacg 6360  
 gggtacctgt ggggagagaa gcaaaaacta ctctatttg tgcatcagat gctaaagcat 6420  
 atgagaaaga agtgcataat gtctggcta cacatgcctg tgtaccacca gaccccaacc 6480  
 cacaagaagt gattttggc aatgtaacag aaaattttaa catgtggaaa aatgacatgg 6540  
 tggatcagat gcaggaagat ataatcagtt tatggatca aagccttaag ccatgtgtaa 6600  
 aattgaccct actctgtgtc actttaaact gtacaatgc aactgttaac tacaataata 6660  
 cctctaaaga catgaaaaat tgctcttct atgttaaccac agaattaaga gataagaaaa 6720  
 agaaagaaaa tgcactttt tatagacttg atatagtacc acttaataat aggaagaatg 6780  
 ggaatattaa caactataga ttaataaaatt gtaataccctc agccataaca caagcctgtc 6840  
 caaaagtctc gtttgaccctt attcctatac attattgtgc tccagctgg tatgcgcctc 6900  
 taaaatgtaa taataagaaa ttcaatggaa taggaccatg cgataatgtc agcacagtac 6960  
 aatgtacaca tggaaattaag ccagtggtat caactcaatt actgttaat ggtagcctag 7020  
 cagaagaaga gataataatt agatctgaaa atctgacaaa caatgtcaaa acaataatag 7080  
 tacatcttaa tgaatctata gagattaaat gtacaagacc tggcaataat acaagaaaga 7140  
 gtgtgagaat aggaccagga caagcattct atgcaacagg agacataataa ggagatataa 7200  
 gacaaggcaca ttgtAACATT agaaaaatg aatggaaatac aactttacaa agggtaagtc 7260  
 aaaaattaca agaactcttc cctaataatgta cagggataaa atttgcacca cactcaggag 7320  
 gggacctaga aattactaca catagctta attgtggagg agaattttt tattgcaata 7380  
 caacagacct gtttaatagt acatacagta atggatcatg cactaatggt acatgcatgt 7440  
 ctaataatac agagcgcattc acactccaat gcagaataaa acaaattata aacatgtggc 7500  
 aggaggtagg acgagcaatg tatgcccctc ccattgcagg aaacataaca tgttagatcaa 7560  
 atattacagg actactatta acacgtatg gaggagataa taataactgaa acagagacat 7620  
 tcagacctgg aggaggagac atgagggaca attggagaag tgaatttat aaatacaagg 7680  
 tggtagaaat taaaccattt ggagttagcac ccactgctgc aaaaaggaga gtgggtggaga 7740  
 gaaaaaaag agcagtagga ataggagctg tggccttgg gttcttggg gcagcaggaa 7800  
 gcactatggg cgcaGcatca ataaacgctga cggtacaggc cagacaatta ttgtctggta 7860  
 tagtgcaca gcaaaataat ttgtcgaggg ctatagaggc gcaacagcat atgttgcac 7920  
 tcacggctcg gggcattaag cagctccagg caagagtctt ggctatagag agataacctac 7980  
 agatcaaca gctccttagga ctgtggggct gctctggaaa actcatctgc accactaatg 8040  
 tgctttggaa ctctagttgg agtaataaaa ctcaaaatgta tatttggat aacatgacct 8100  
 ggtgcagtg ggataggaa attagtaatt acacaaacac aatatacagg ttgcttgaag 8160  
 actcgcaag ccagcaggaa agaaatgaaa aagatttact agcattggac aggtggaaca 8220  
 atctgtggaa ttgggtttagc ataacaaattt ggctgtggta tataaaaata ttcataatga 8280  
 tagtaggagg ctgtataggt ttaagaataa ttttgctgt gctctctta gtaaatagag 8340  
 ttaggcaggg atactcaccc ttgtcattgc agacccttat cccaaacccg aggggacccg 8400  
 acaggctcgg aggaatcgaa gaagaagggtg gagagcaaga cagcagcaga tccattcgat 8460  
 tagtgagcgg attcttgaca ctggccttgg acgacactacg aagcctgtgc ctcttctgt 8520  
 accaccgatt gagagacttc atattaattt tagtgagacg agtggaaactt ctggacaca 8580  
 gtagtctcg gggactgcag aggggggtggg gaacccttaa gtatttgggg agtctgtgc 8640  
 aatattgggg tctagagttt aaaaagagtg ctattaatct gcttgataact atagcaatag 8700  
 cagtagctga aggaacagat aggattctag aattcataca aaacctttt agaggtatcc 8760  
 gcaacgtacc tagaagaata agacagggtc tcgaagcgc tttgcaataa aatggggggc 8820  
 aagtggtaaa aaagcagtat aattggatgg cctgaagttt gagaagaat cagacgaact 8880  
 aggtcagcag cagaggagt aggtcagcg tctcaagact tagagaaaca tggggcactt 8940  
 acaaccagca acacagcccc caacaatgct gcttgcgcct ggctggaaagc gcaagaggag 9000  
 gaaggagaag taggcttcc agtcagacct caggtacctt taagaccaat gacttataaa 9060  
 gcagcaatag atctcagctt cttttaaaa gaaaaggggg gactggaaagg gttaatttac 9120  
 tccaagaaaa ggcaagagat ctgtgatTT tgggttata acacacaagg cttctccct 9180  
 gatggcaaa actacacacc gggaccaggc gtcagattc cactgacctt tggatggtag 9240  
 ttcaagctag agccagtcga tccaaggggaa gttagaagagg ccaatgaagg agaaaaacaac 9300  
 tggtagttt accctatgag ccagcatggc atggaggatg aagacagaga agtattaaga 9360  
 tggtagttt acagtagctt agcacgcaga cacatggccc gcgagctaca tccggaggat 9420  
 tacaagact gctgacacag aaggggacttt ccgctggac tttccactgg ggcgttccag 9480  
 gaggtgttgt ctggcgggaa cagggagtg gtcagccctg agatgctgca tataagcagc 9540  
 tgctttcgc ctgtacttggg tctctcttagg tagaccagat ctgagcccg gagctctctg 9600

gctatctagg gaacccactg cttaagcctc aataaagctt gccttgagtgc ttgtggtag 9660  
tgtgtccccg tctgttgtgt gactctggta actagagatc ctcagaccca cttgtggtag 9720  
tgtggaaaat ctctagca 9738

<210> 46  
<211> 4  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: spacer

<400> 46  
Gly Gly Gly Ser  
1